



Universidade Federal do Amazonas
Instituto de Computação
Programa de Pós-Graduação em Informática
Grupo de Interesse em Sistemas Embarcados

**EXTRAÇÃO DESCENTRALIZADA DE CONHECIMENTO ASSOCIATIVO
PARA INTERNET DAS COISAS**

Prof. Dr. Raimundo Barreto
Márcio André da Costa Alencar

Manaus, 25 de março de 2019

AGENDA



INTRODUÇÃO

I

Contextualização

PROBLEMÁTICA

II

Definição do problema e objetivos

MÉTODO PROPOSTO

III

Descrição dos componentes e arquitetura geral

EXPERIMENTOS

IV

Métodologia de avaliação

RESULTADOS

V

Apresentação dos resultados

CONCLUSÃO

VI

Considerações Finais

INTRODUÇÃO



Contextualização

I

- O que é “*Internet of Things*” ?

II

III

IV

V

VI

INTRODUÇÃO

Contextualização

- I - O que é “*Internet of Things*” ? Não há uma definição universal mas, simplificando, é o cenário onde as “coisas” estão conectadas!



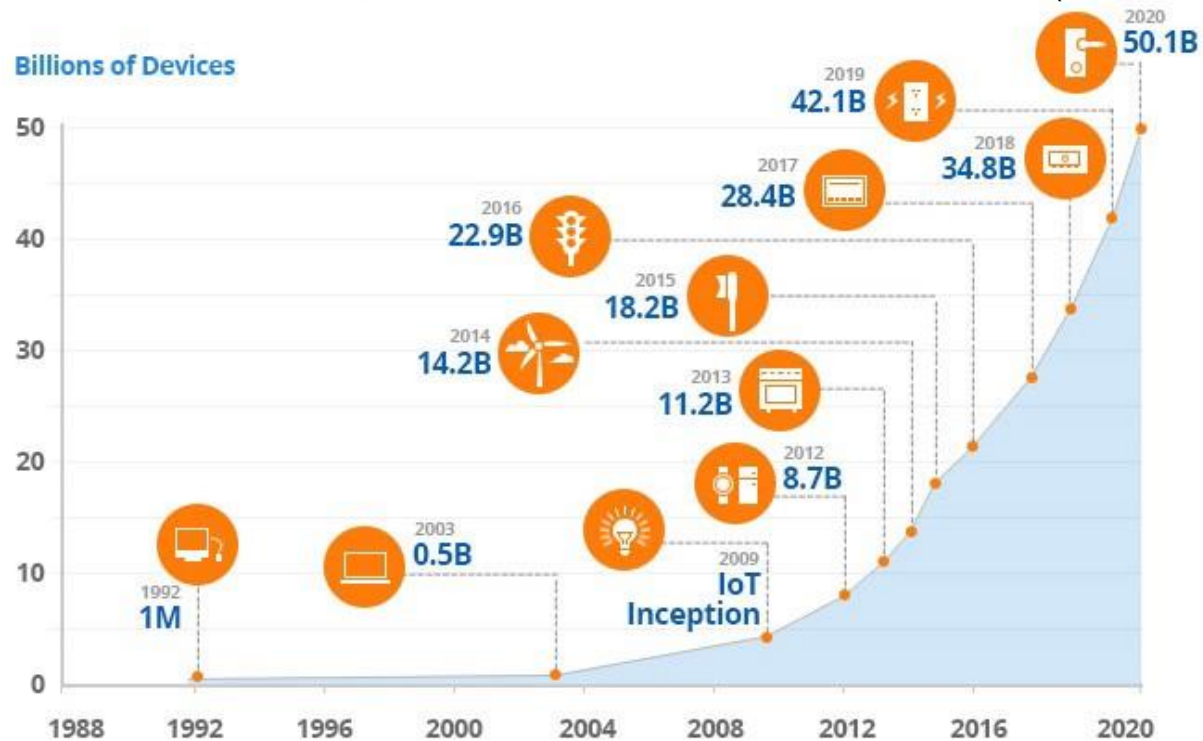
INTRODUÇÃO



UFAM

Contextualização

- Há aproximadamente 1 bilhão de dispositivos conectados diariamente. Até 2020 haverá mais de 50 bilhões (IDC - 2011)



INTRODUÇÃO

Contextualização

- Muitos deles estarão em nossas casas e trabalhos



INTRODUÇÃO

Contextualização

- Muitos deles estarão em nossas casas e trabalhos
- Estar conectado não é o suficiente, é preciso inteligência!



PROBLEMÁTICA



Como prover inteligência aos dispositivos IoT?

I

II

III

IV

V

VI

PROBLEMÁTICA



Como prover inteligência aos dispositivos IoT?

I

É necessário que estes sejam capazes de identificar atividades, padrões, tendências e/ou correlações implícitas entre as atividades dos dispositivos que compõem o ambiente inteligente

II

III

IV

V

VI

PROBLEMÁTICA



Como prover inteligência aos dispositivos IoT?

I

É necessário que estes sejam capazes de identificar atividades, padrões, tendências e/ou correlações implícitas entre as atividades dos dispositivos que compõem o ambiente inteligente

III

Mas para isso é necessário considerar os seguintes pontos:

IV

V

VI

PROBLEMÁTICA



Como prover inteligência aos dispositivos IoT?

I

É necessário que estes sejam capazes de identificar atividades, padrões, tendências e/ou correlações implícitas entre as atividades dos dispositivos que compõem o ambiente inteligente.

II

III

Mas para isso é necessário considerar os seguintes pontos:

IV

- **Arquitetura:** Centralizada / Descentralizada

- **Tipo de inteligência:** Que informações extrair?

V

- **Limitações dos dispositivos:** Muito/Poucos Recursos

- **Custo/Benefício:** Soluções caras ou econômicas

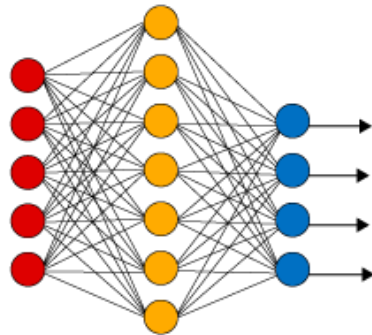
VI

PROBLEMÁTICA

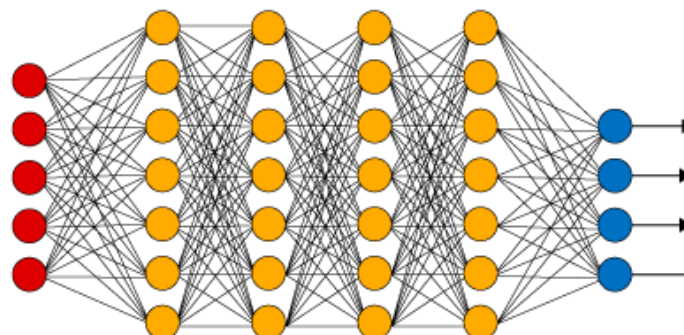
Estado da arte em reconhecimento de padrões

- Redes Neurais (Aprendizagem profunda)

Simple Neural Network



Deep Learning Neural Network



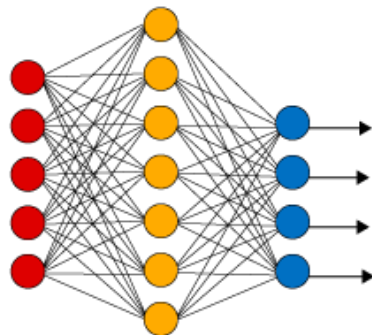
● Input Layer ● Hidden Layer ● Output Layer

PROBLEMÁTICA

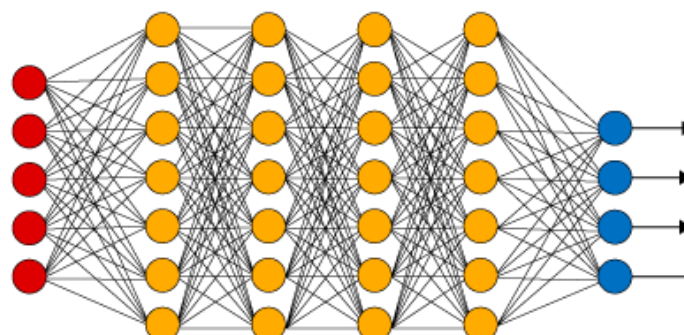
Estado da arte em reconhecimento de padrões

- Redes Neurais (Aprendizagem profunda)
- Alto custo computacional (armazenamento/processamento)
(CHEN *et al.*, 2015; ROSE *et al.*, 2015; MOONS e VERHELST, 2017)

Simple Neural Network



Deep Learning Neural Network

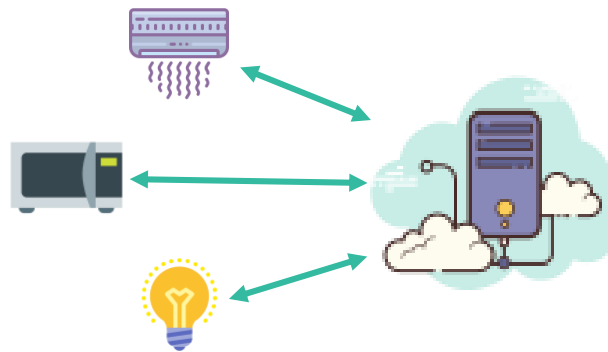


● Input Layer ● Hidden Layer ● Output Layer

PROBLEMÁTICA

Estado da arte em reconhecimento de padrões

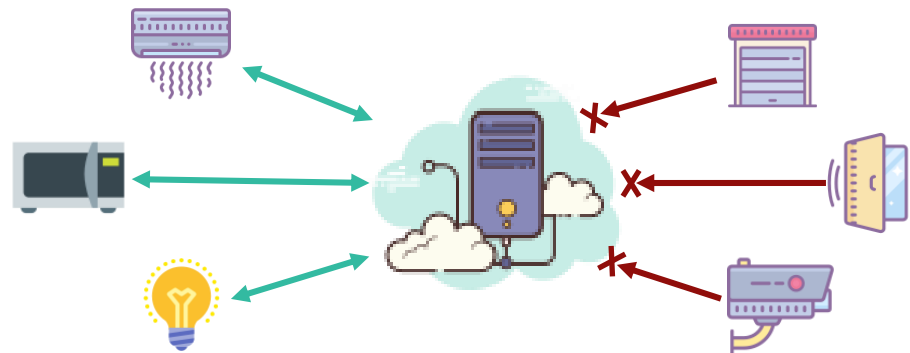
- I - Redes Neurais (Aprendizagem profunda)
- II - Alto custo computacional (armazenamento/processamento)
(CHEN *et al.*, 2015; ROSE *et al.*, 2015; MOONS e VERHELST, 2017)
- III - Centralização / Dependência



PROBLEMÁTICA

Estado da arte em reconhecimento de padrões

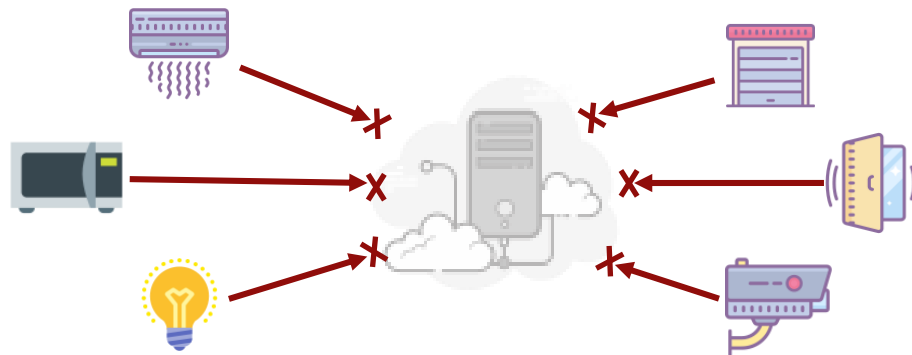
- I - Redes Neurais (Aprendizagem profunda)
- II - Alto custo computacional (armazenamento/processamento)
(CHEN *et al.*, 2015; ROSE *et al.*, 2015; MOONS e VERHELST, 2017)
- III - Centralização / Dependência



PROBLEMÁTICA

Estado da arte em reconhecimento de padrões

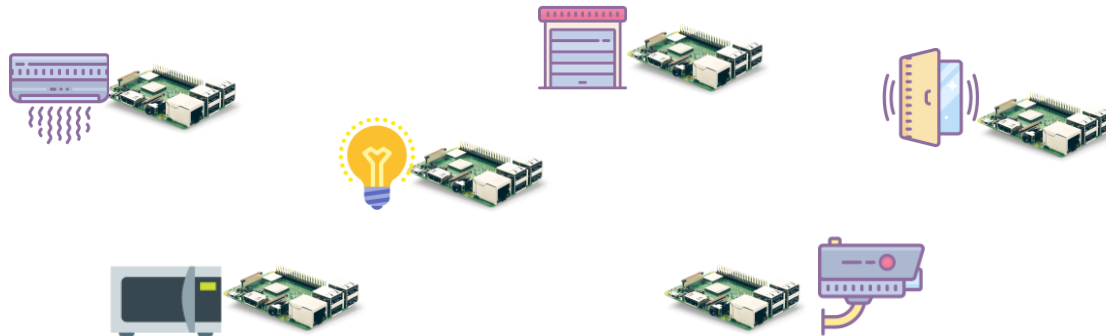
- I - Redes Neurais (Aprendizagem profunda)
- II - Alto custo computacional (armazenamento/processamento)
(CHEN *et al.*, 2015; ROSE *et al.*, 2015; MOONS e VERHELST, 2017)
- III - Centralização / Dependência



PROBLEMÁTICA

Estado da arte em reconhecimento de padrões

- I - Redes Neurais (Aprendizagem profunda)
- II - Alto custo computacional (armazenamento/processamento)
(CHEN *et al.*, 2015; ROSE *et al.*, 2015; MOONS e VERHELST, 2017)
- III - Centralização / Dependência
- IV - Dispositivos com muitos recursos (alto custo/dispercício)



PROBLEMÁTICA



Como prover inteligência aos dispositivos IoT...

I Considerando:

- Ambiente descentralizado
- Dispositivos de baixo custos:
 - Limitações de processamento e armazenamento
- Extração de conhecimento deste ambiente

II

III

IV

V

VI

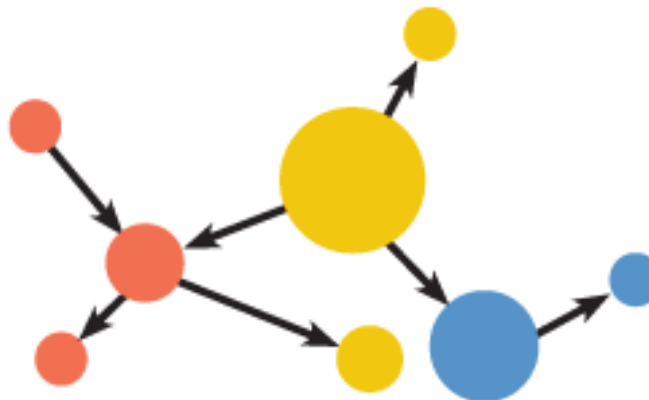
PROBLEMÁTICA



Mineração de dados!

I Uma estratégia para contornar tais limitações é através da
II mineração de dados, especialmente a **análise associativa**, que
III buscam identificar padrões similares em um conjunto dados de
IV modo que satisfaça critérios mínimos de suporte e confiabilidade

(TAN et al. 2006; CHEN et al. 2015; LI et al. 2018;
NAZERFARD 2018; KIREEV et al. 2019)

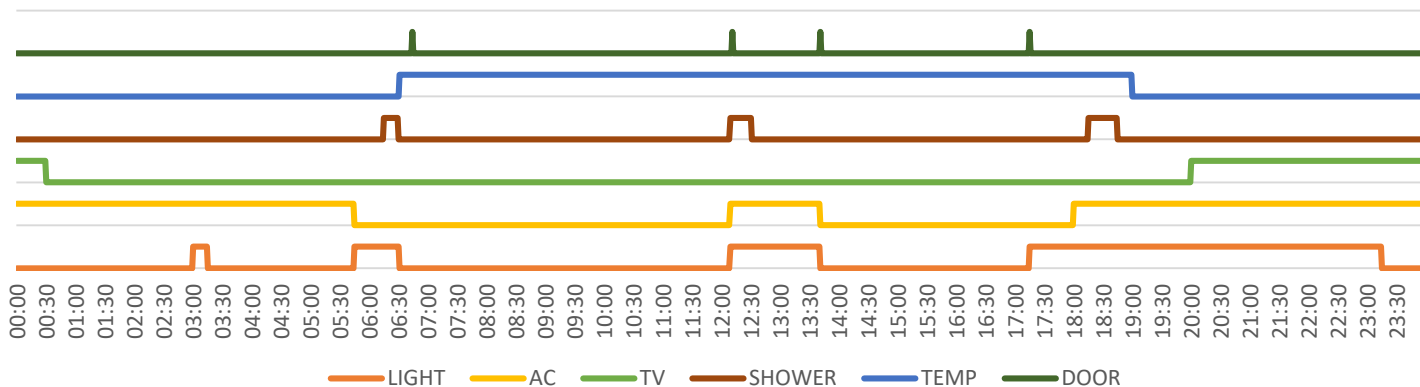


PROBLEMÁTICA

Mineração de dados!

Embora cada dispositivo tenha um padrão singular de uso

(HEIERMAN *et al.*, 2003; CHEN *et al.*, 2012; GONZALES e AMFT, 2015;
PAL *et al.*, 2017; KIREEV *et al.* 2018; NAZERFARD 2018;).

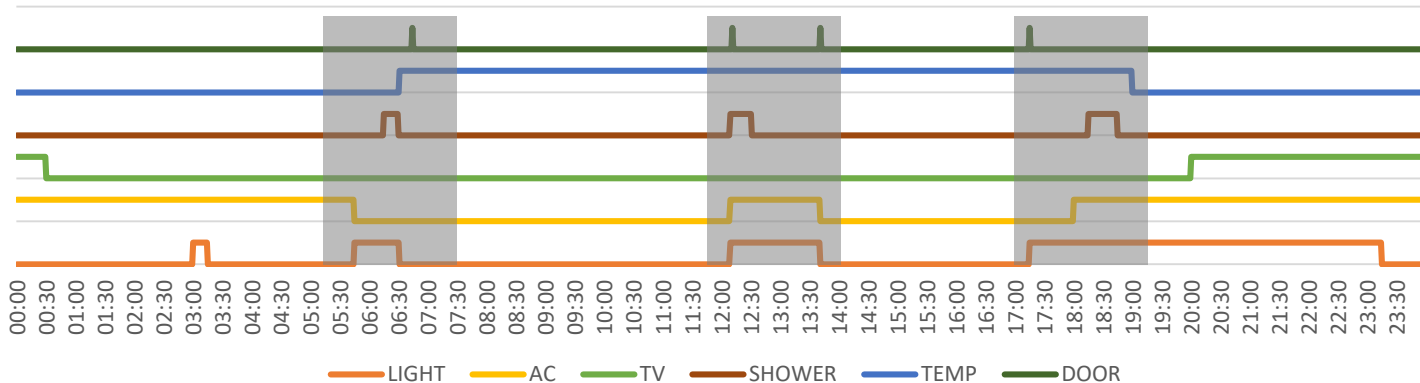


PROBLEMÁTICA

Mineração de dados!

Embora cada dispositivo tenha um padrão singular de uso, alguns deles possuem similaridades que podem ser exploradas para correlacioná-los

(HEIERMAN *et al.*, 2003; CHEN *et al.*, 2012; GONZALES e AMFT, 2015; PAL *et al.*, 2017; KIREEV *et al.* 2018; NAZERFARD 2018;).



MÉTODO PROPOSTO



Informações Preliminares

I

É importante assimilar bem os componentes e especificações do método para melhor compreensão da visão geral da arquitetura

II

III

IV

V

VI

MÉTODO PROPOSTO



Informações Preliminares

I

É importante assimilar bem os componentes e especificações do método para melhor compreensão da visão geral da arquitetura

II

Logo, tenha em mente que:

- **Arquitetura:** Descentralizada (tudo embarcado)
- **Tipo de inteligência:** Extração de correlações
- **Dispositivos:** Baixo custo (Poucos Recursos). Ex. ESP8266-01
- **Desafio:** Obter um conhecimento global baseado em decisões locais

V

VI

Comportamento dos dispositivos

I

Cada dispositivo controla um **único objeto** e deve declarar **dois conjuntos** bem definidos (discretos): Conjunto de **estados** $S = \{s_1, s_2, \dots, s_k\}$ que representa suas interações com o ambiente e o conjunto de **ações** $A = \{a_1, a_2, \dots, a_i\}$ que permite transitar entre os $|S|$ estados dos dispositivo

II

III

IV

V

VI

MÉTODO PROPOSTO

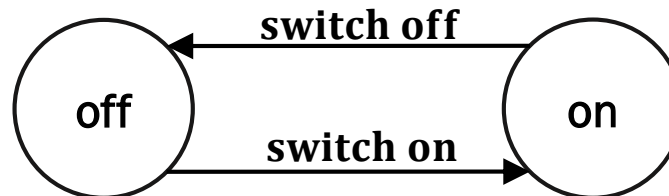


Comportamento dos dispositivos

I Cada dispositivo controla um **único objeto** e deve declarar **dois conjuntos** bem definidos (discretos): Conjunto de **estados** $S = \{s_1, s_2, \dots, s_k\}$ que representa suas interações com o ambiente e o conjunto de **ações** $A = \{a_1, a_2, \dots, a_i\}$ que permite transitar entre os $|S|$ estados dos dispositivo. Ex.: Dispositivos binário (lâmpada)

$$S = \{off, on\}$$

$$A = \{switch\ off, switch\ on\}$$



MÉTODO PROPOSTO



Armazenamento embarcado

I

Seja $T = \{t_1, t_2, \dots, t_j\}$ um conjunto finito de intervalos de tempos discretos (*slots*), é possível definir uma base de dados embarcada em forma de matriz $M_{ij} = A \times T$ onde cada elemento $c_{xy} \in M_{ij}$ é um contador para a ação $a_x \in A$ no slot $t_y \in T$

II

III

IV

V

VI

MÉTODO PROPOSTO



Armazenamento embarcado

I
II
III
Seja $T = \{t_1, t_2, \dots, t_j\}$ um conjunto finito de intervalos de tempos discretos (*slots*), é possível definir uma base de dados embarcada em forma de matriz $M_{ij} = A \times T$ onde cada elemento $c_{xy} \in M_{ij}$ é um contador para a ação $a_x \in A$ no slot $t_y \in T$

IV
V
VI

$$M_{ij} = \begin{bmatrix} c_{11} & \cdots & c_{1j} \\ \vdots & \ddots & \vdots \\ c_{i1} & \cdots & c_{ij} \end{bmatrix}$$

i : num de elementos em A ($|A|$);
 j : num de elementos em T ($|T|$);
 c_{xy} : contador para cada ação a_x no slot t_y ;

MÉTODO PROPOSTO



Armazenamento embarcado

I
II
III
Seja $T = \{t_1, t_2, \dots, t_j\}$ um conjunto finito de intervalos de tempos discretos (*slots*), é possível definir uma base de dados embarcada em forma de matriz $M_{ij} = A \times T$ onde cada elemento $c_{xy} \in M_{ij}$ é um contador para a ação $a_x \in A$ no slot $t_y \in T$

IV
V
VI
$$M_{ij} = \begin{bmatrix} c_{11} & \dots & c_{1j} \\ \vdots & \ddots & \vdots \\ c_{i1} & \dots & c_{ij} \end{bmatrix}$$

 i : num de elementos em A ($|A|$);
 j : num de elementos em T ($|T|$);
 c_{xy} : contador para cada ação a_x no slot t_y ;

$M_{ij} =$

	00:00	01:00	02:00	...	21:00	22:00	23:00
switch off	10	2	1	...	26	0	12
switch on	2	2	0	...	34	5	7

MÉTODO PROPOSTO



Padrão de ações

I

A partir da base de dados embarcada é possível **extrair o padrão de ações** do dispositivo. Pra isso é **necessário** realizar o tratamento de dados através de uma **transformação logarítmica** em cada contador para **remover outliers** e **reduzir peso de informações antigas**

II

III

IV

V

VI

MÉTODO PROPOSTO



Padrão de ações

I

A partir da base de dados embarcada é possível **extrair o padrão de ações** do dispositivo. Pra isso é **necessário** realizar o tratamento de dados através de uma **transformação logarítmica** em cada contador para **remover outliers** e **reduzir peso** de informações antigas

II

III

IV

V

VI

$$\forall c_{xy} \in M_{ij} \text{ temos } c_{xy} = \begin{cases} \log \frac{c_{xy}}{|A|} & \text{se } |A| \geq c_{xy} \\ 0 & \text{se } |A| < c_{xy} \end{cases}$$

Evita valores de inferiores à 1 para não gerar resultados negativos em transformações futuras

MÉTODO PROPOSTO



Padrão de ações

M_{ij}

	00:00	01:00	02:00	...	21:00	22:00	23:00
switch off	10	2	1	...	26	0	12
switch on	2	2	0	...	34	5	7

MÉTODO PROPOSTO



UFAM

Padrão de ações

I

II

III

IV

V

VI

M_{ij}

	00:00	01:00	02:00	...	21:00	22:00	23:00
switch off	10	2	1	...	26	0	12
switch on	2	2	0	...	34	5	7



M'_{ij}

	00:00	01:00	02:00	...	21:00	22:00	23:00
switch off	3,32	1	0	...	4,70	0	3,58
switch on	1	1	0	...	5,09	2,32	2,81

MÉTODO PROPOSTO

Padrão de ações

I

II

III

IV

V

VI

M_{ij}

	00:00	01:00	02:00	...	21:00	22:00	23:00
switch off	10	2	1	...	26	0	12
switch on	2	2	0	...	34	5	7



M'_{ij}

	00:00	01:00	02:00	...	21:00	22:00	23:00
switch off	<u>3,32</u>	1	0	...	4,70	0	<u>3,58</u>
switch on	1	1	0	...	<u>5,09</u>	<u>2,32</u>	2,81



Padrão de
ações (P)

00:00	01:00	02:00	...	21:00	22:00	23:00
SWITCH OFF	-	-	...	SWITCH ON	SWITCH ON	SWITCH OFF

MÉTODO PROPOSTO



UFAM

Minerando correlações

I

II

III

IV

V

VI

Padrão do
Dispositivo 1

00:00	01:00	02:00	...	21:00	22:00	23:00
SWITCH OFF	-	-	...	SWITCH ON	SWITCH ON	SWITCH OFF

MÉTODO PROPOSTO



Minerando correlações

I

II

III

IV

V

VI

Padrão do
Dispositivo 1

00:00	01:00	02:00	...	21:00	22:00	23:00
SWITCH OFF	-	-	...	SWITCH ON	SWITCH ON	SWITCH OFF

Padrão do
Dispositivo 2

00:00	01:00	02:00	...	21:00	22:00	23:00
CLOSE DOOR	CLOSE DOOR	-	...	OPEN DOOR	CLOSE DOOR	CLOSE DOOR

MÉTODO PROPOSTO



Minerando correlações

I

II

III

IV

V

VI

Padrão do Dispositivo 1

Padrão do Dispositivo 2

Base de Transações (D)

00:00	01:00	02:00	...	21:00	22:00	23:00
SWITCH OFF	-	-	...	SWITCH ON	SWITCH ON	SWITCH OFF
00:00	01:00	02:00	...	21:00	22:00	23:00
CLOSE DOOR	CLOSE DOOR	-	...	OPEN DOOR	CLOSE DOOR	CLOSE DOOR
00:00	01:00	02:00	...	21:00	22:00	23:00
SWITCH OFF	-	-	...	SWITCH ON	SWITCH ON	SWITCH OFF
CLOSE DOOR	CLOSE DOOR	-	...	OPEN DOOR	CLOSE DOOR	CLOSE DOOR

MÉTODO PROPOSTO

Minerando correlações

I

Padrão do Dispositivo 1

00:00

01:00

02:00

...

21:00

22:00

23:00

SWITCH OFF

-

-

...

SWITCH ON

SWITCH ON

SWITCH OFF

III

Padrão do Dispositivo 2

00:00

01:00

02:00

...

21:00

22:00

23:00

CLOSE DOOR

CLOSE DOOR

-

...

OPEN DOOR

CLOSE DOOR

CLOSE DOOR

IV

Base de Transações (D)

00:00

01:00

02:00

...

21:00

22:00

23:00

SWITCH OFF

-

-

...

SWITCH ON

SWITCH ON

SWITCH OFF

CLOSE DOOR

CLOSE DOOR

-

...

OPEN DOOR

CLOSE DOOR

CLOSE DOOR

VI

Combinações das ações (Regras)

SWITCH OFF ⇒ CLOSE DOOR
 SWITCH OFF ⇒ OPEN DOOR
 SWITCH ON ⇒ OPEN DOOR
 SWITCH ON ⇒ CLOSE DOOR

MÉTODO PROPOSTO



Minerando correlações

I

II

III

IV

V

VI

Padrão do Dispositivo 1

Padrão do Dispositivo 2

Base de Transações (D)

Combinações das ações (Regras)

00:00	01:00	02:00	...	21:00	22:00	23:00
SWITCH OFF	-	-	...	SWITCH ON	SWITCH ON	SWITCH OFF
00:00	01:00	02:00	...	21:00	22:00	23:00
CLOSE DOOR	CLOSE DOOR	-	...	OPEN DOOR	CLOSE DOOR	CLOSE DOOR
00:00	01:00	02:00	...	21:00	22:00	23:00
SWITCH OFF	-	-	...	SWITCH ON	SWITCH ON	SWITCH OFF
CLOSE DOOR	CLOSE DOOR	-	...	OPEN DOOR	CLOSE DOOR	CLOSE DOOR

SWITCH OFF ⇒ CLOSE DOOR
 SWITCH OFF ⇒ OPEN DOOR
 SWITCH ON ⇒ OPEN DOOR
 SWITCH ON ⇒ CLOSE DOOR

Quais regras são
Regras mais interessantes!?

MÉTODO PROPOSTO



Minerando correlações

I

Uma regra de associação, correlaciona a **Ação X** do **Dispositivo 01** à **Ação Y** do **Dispositivo 02** e expressa sua relevância baseada nas seguintes métricas:

II

III

IV

V

VI

MÉTODO PROPOSTO



Minerando correlações

Uma regra de associação, correlaciona a **Ação X** do **Dispositivo 01** à **Ação Y** do **Dispositivo 02** e expressa sua relevância baseada nas seguintes métricas:

Regra $X \Rightarrow Y$

- $Support = \frac{freq(X, Y)}{|D|}$
- $Lift = \frac{freq(X, Y) \times |D|}{freq(X) \times freq(Y)}$
- $Confidence = \frac{freq(X, Y)}{freq(X)}$

MÉTODO PROPOSTO



Minerando correlações

Uma regra de associação, correlaciona a **Ação X** do **Dispositivo 01** à **Ação Y** do **Dispositivo 02** e expressa sua relevância baseada nas seguintes métricas:

Regra $X \Rightarrow Y$

- $Support = \frac{freq(X, Y)}{|D|}$
- $Lift = \frac{freq(X, Y) \times |D|}{freq(X) \times freq(Y)}$
- $Confidence = \frac{freq(X, Y)}{freq(X)}$

MÉTODO PROPOSTO



Minerando correlações

Uma regra de associação, correlaciona a **Ação X** do **Dispositivo 01** à **Ação Y** do **Dispositivo 02** e expressa sua relevância baseada nas seguintes métricas:

Regra $X \Rightarrow Y$

- $Support = \frac{freq(X, Y)}{|D|}$
- $Lift = \frac{freq(X, Y) \times |D|}{freq(X) \times freq(Y)}$
- $Confidence = \frac{freq(X, Y)}{freq(X)}$

Diagram illustrating the relationship between the metrics and the dataset size $|D|$:

- A red arrow points from the $|D|$ in the Support formula to $|D| = |T|$.
- A red arrow points from the $|D|$ in the Lift formula to $|D| = |T|$.

Esta abordagem define que $|D| = |T|$ para comparar, de forma justa, regras de bases de transações com tamanhos diferentes

MÉTODO PROPOSTO



Minerando correlações

I

Para obter as regras mais relevantes para suas ações, o dispositivo deve combinar seu padrão de ação com o de cada dispositivo na rede, extrair as regras e comparar com as já obtidas até que reste apenas as mais relevantes

II

III

IV

V

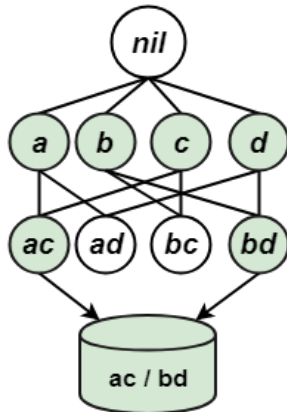
VI

MÉTODO PROPOSTO

Minerando correlações

I Para obter as regras mais relevantes para suas ações, o dispositivo
II deve combinar seu padrão de ação com o de cada dispositivo na
III rede, extrair as regras e comparar com as já obtidas até que reste
IV apenas as mais relevantes

Iteração 01
D01 x D02



Armazenar as
regras ac e bd

D01 : AÇÕES (a, b)
D02 : AÇÕES (c, d)

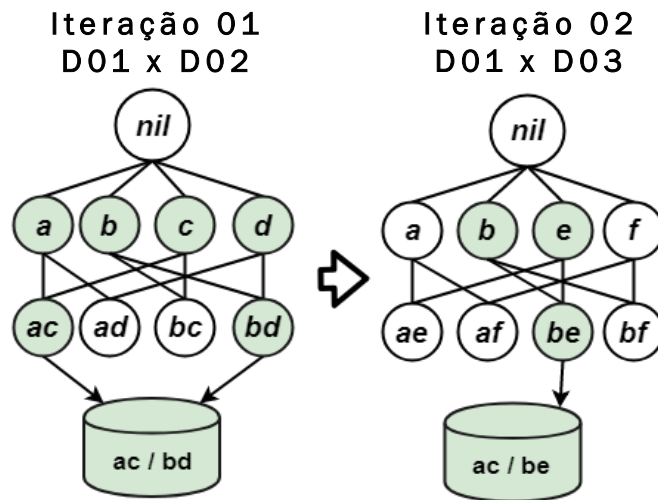
MÉTODO PROPOSTO



UFAM

Minerando correlações

Para obter as regras mais relevantes para suas ações, o dispositivo deve combinar seu padrão de ação com o de cada dispositivo na rede, extrair as regras e comparar com as já obtidas até que reste apenas as mais relevantes.



Armazenar as
regras ac e bd

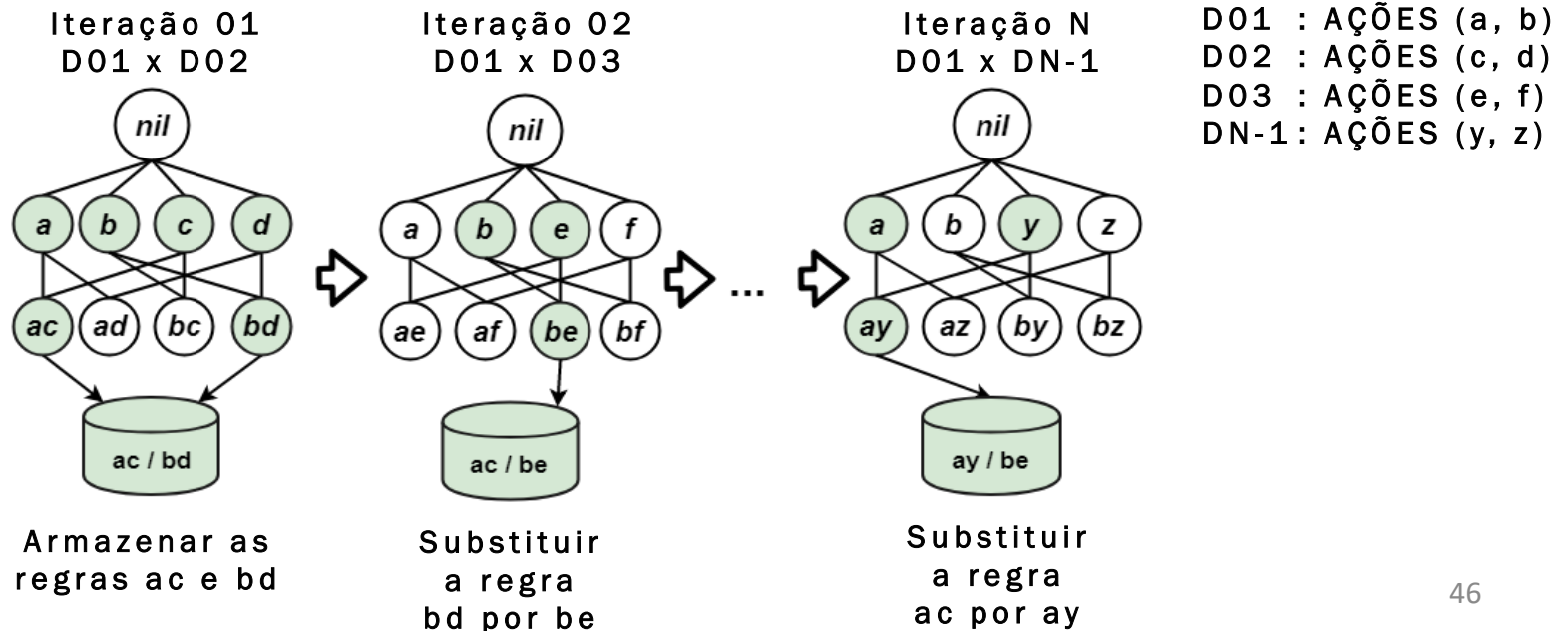
Substituir
a regra
bd por be

D01 : AÇÕES (a, b)
D02 : AÇÕES (c, d)
D03 : AÇÕES (e, f)

MÉTODO PROPOSTO

Minerando correlações

Para obter as regras mais relevantes para suas ações, o dispositivo deve combinar seu padrão de ação com o de cada dispositivo na rede, extrair as regras e comparar com as já obtidas até que reste apenas as mais relevantes.



MÉTODO PROPOSTO



Usando as regras

I

Baseado na regra, o Dispositivo 01 pode solicitar que o Dispositivo 02 mude de estado para satisfazer a regra de correlação no *slot* atual

II

III

IV

V

VI

MÉTODO PROPOSTO

Usando as regras

I

Baseado na regra, o Dispositivo 01 pode solicitar que o Dispositivo 02 mude de estado para satisfazer a regra de correlação no *slot* atual

II

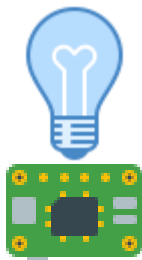
III

REGRAS NO DISPOSITIVO 01
SWITCH ON \Rightarrow CLOSE DOOR

IV

V

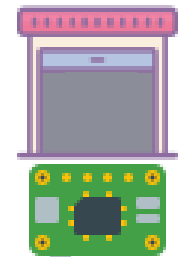
VI



DISPOSITIVO 01

00:00	SWITCH OFF
01:00	-
02:00	-
...	...
21:00	SWITCH ON
22:00	SWITCH ON
23:00	SWITCH OFF

00:00	CLOSE DOOR
01:00	CLOSE DOOR
02:00	-
...	...
21:00	OPEN DOOR
22:00	CLOSE DOOR
23:00	CLOSE DOOR



DISPOSITIVO 02

MÉTODO PROPOSTO

Usando as regras

I

Baseado na regra, o Dispositivo 01 pode solicitar que o Dispositivo 02 mude de estado para satisfazer a regra de correlação no *slot* atual

II

III

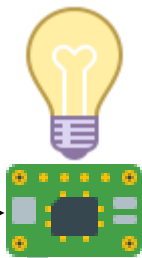
REGRAS NO DISPOSITIVO 01
SWITCH ON \Rightarrow CLOSE DOOR

IV

V

VI

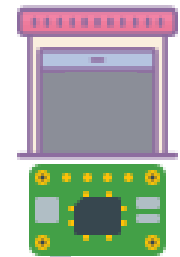
SWITCH ON



DISPOSITIVO 01

00:00	SWITCH OFF
01:00	-
02:00	-
...	...
21:00	SWITCH ON
22:00	SWITCH ON
23:00	SWITCH OFF

00:00	CLOSE DOOR
01:00	CLOSE DOOR
02:00	-
...	...
21:00	OPEN DOOR
22:00	CLOSE DOOR
23:00	CLOSE DOOR



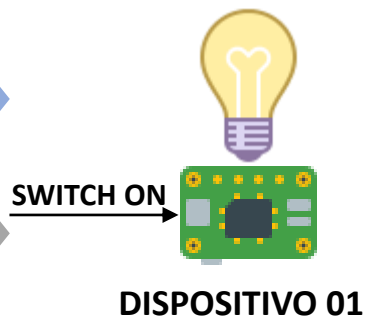
DISPOSITIVO 02

MÉTODO PROPOSTO

Usando as regras

I Baseado na regra, o Dispositivo 01 pode solicitar que o Dispositivo 02 mude de estado para satisfazer a regra de correlação no slot atual

REGRAS NO DISPOSITIVO 01
SWITCH ON \Rightarrow CLOSE DOOR

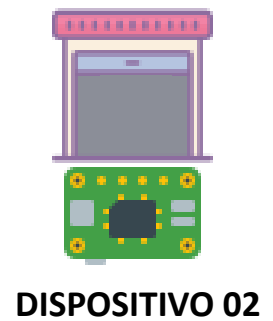


00:00	SWITCH OFF	×
01:00	-	×
02:00	-	×
...	...	
21:00	SWITCH ON	
22:00	SWITCH ON	
23:00	SWITCH OFF	×

CLOSE DOOR

CLOSE DOOR

00:00	CLOSE DOOR
01:00	CLOSE DOOR
02:00	-
...	...
21:00	OPEN DOOR
22:00	CLOSE DOOR
23:00	CLOSE DOOR



MÉTODO PROPOSTO

Usando as regras

I

Baseado na regra, o Dispositivo 01 pode solicitar que o Dispositivo 02 mude de estado para satisfazer a regra de correlação no *slot* atual

II

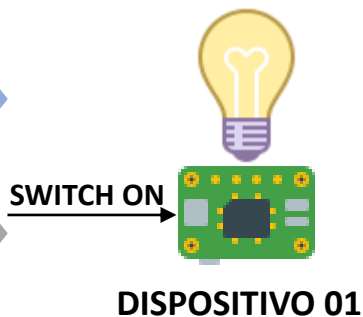
III

IV

V

VI

REGRAS NO DISPOSITIVO 01 SWITCH ON \Rightarrow CLOSE DOOR

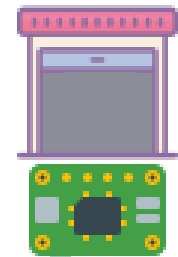


00:00	SWITCH OFF
01:00	-
02:00	-
...	...
21:00	SWITCH ON
22:00	SWITCH ON
23:00	SWITCH OFF

CLOSE DOOR

00:00	CLOSE DOOR
01:00	CLOSE DOOR
02:00	-
...	...
21:00	OPEN DOOR
22:00	CLOSE DOOR
23:00	CLOSE DOOR

✗



DISPOSITIVO 02

MÉTODO PROPOSTO



UFAM

Usando as regras

I

Baseado na regra, o Dispositivo 01 pode solicitar que o Dispositivo 02 mude de estado para satisfazer a regra de correlação no *slot* atual

II

III

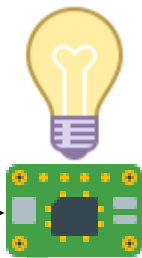
REGRAS NO DISPOSITIVO 01
SWITCH ON \Rightarrow CLOSE DOOR

IV

V

VI

SWITCH ON

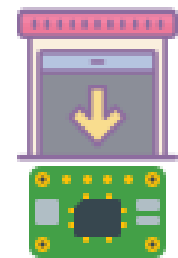


DISPOSITIVO 01

00:00	SWITCH OFF
01:00	-
02:00	-
...	...
21:00	SWITCH ON
22:00	SWITCH ON
23:00	SWITCH OFF

CLOSE DOOR

00:00	CLOSE DOOR
01:00	CLOSE DOOR
02:00	-
...	...
21:00	OPEN DOOR
22:00	CLOSE DOOR
23:00	CLOSE DOOR



DISPOSITIVO 02

MÉTODO PROPOSTO



UFAM

Visão geral

I

Dispositivos seguem especificações da *Web Thing Model* e compartilham um mesmo grupo de *multicast*

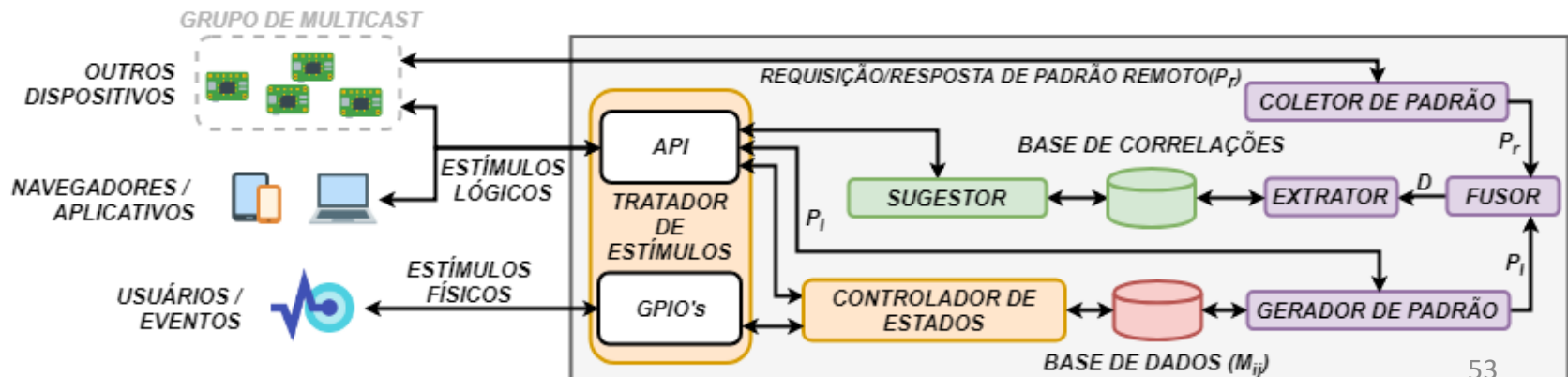
II

III

IV

V

VI

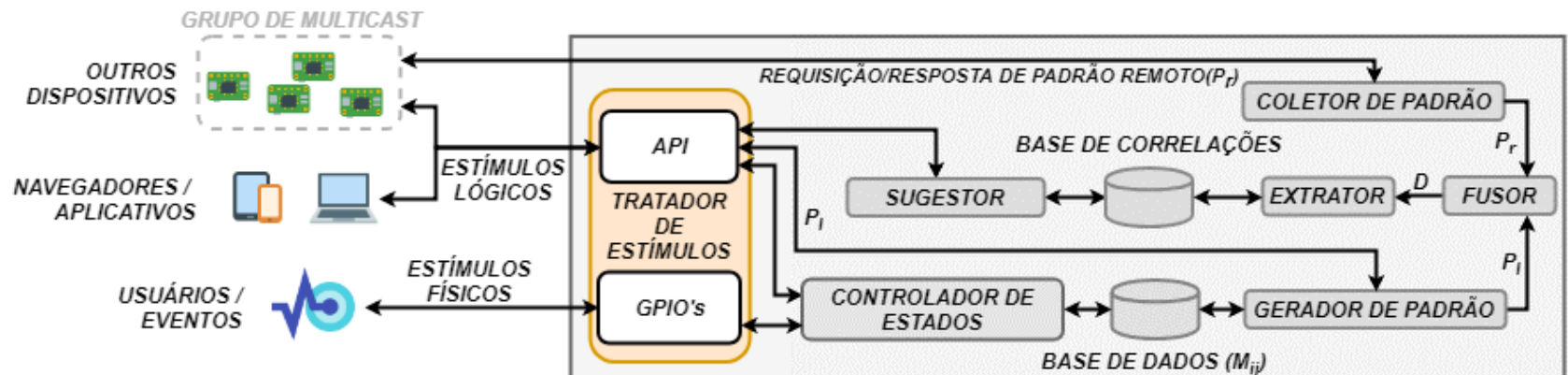


MÉTODO PROPOSTO

Visão geral

I Dispositivos seguem especificações da *Web Thing Model* e compartilham um mesmo grupo de *multicast*

- II
- III
- **Tratador de estímulos:** Interage com o ambiente. Identifica entradas e gera saídas do dispositivo, Disponibiliza as ferramentas para controlar o estado do dispositivo, acessar o padrão de ações e gerenciar as regras de correlação, além de outras personalizações
- IV
- V
- VI



MÉTODO PROPOSTO

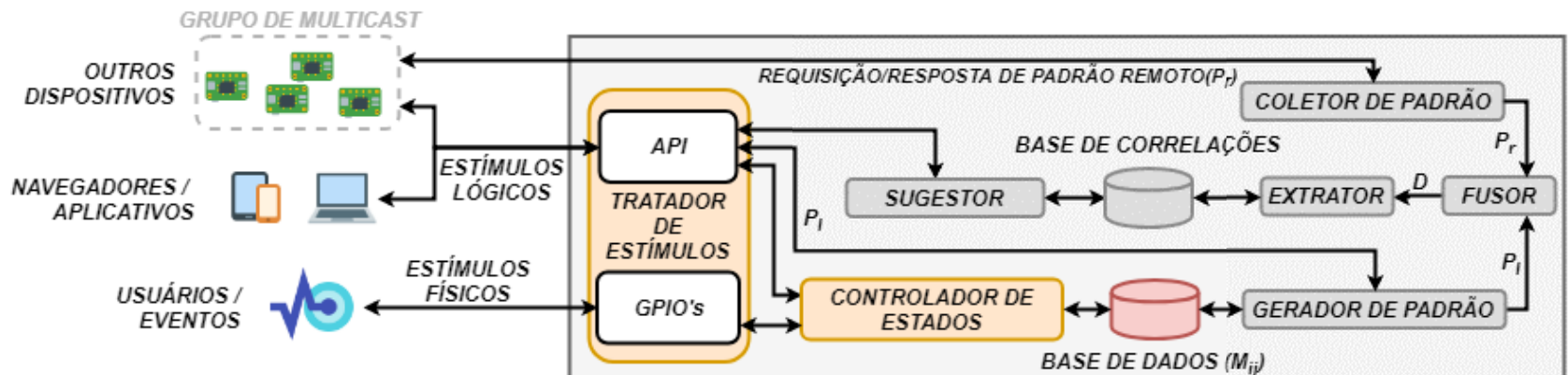


UFAM

Visão geral

I Dispositivos seguem especificações da *Web Thing Model* e compartilham um mesmo grupo de *multicast*

- II
- III
- IV **Controlador de estados:** Realiza a mudança de estado se o estímulo for uma ação válida. Analisa a ação mais provável de ocorrer no *slot* antes de incrementar o contador. Informa ao Tratador de Estímulos se deve ou não disparar uma requisição para o dispositivo remoto
- V
- VI



MÉTODO PROPOSTO

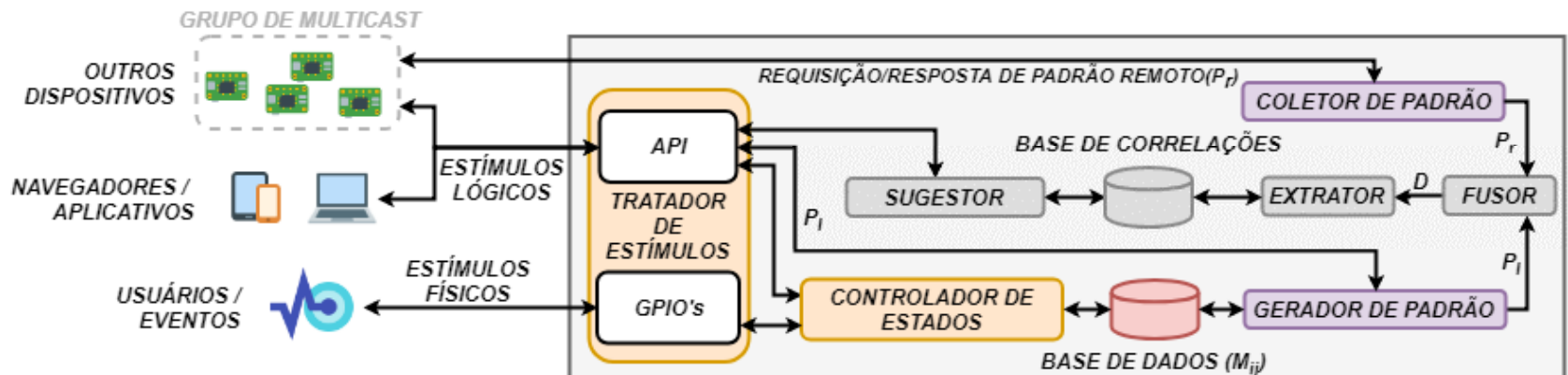


UFAM

Visão geral

I Dispositivos seguem especificações da *Web Thing Model* e compartilham um mesmo grupo de *multicast*

- III • **Coletor de padrão:** Identifica os demais dispositivos na rede por meio do grupo de *multicast* e coleta o padrão remoto dos dispositivos



MÉTODO PROPOSTO

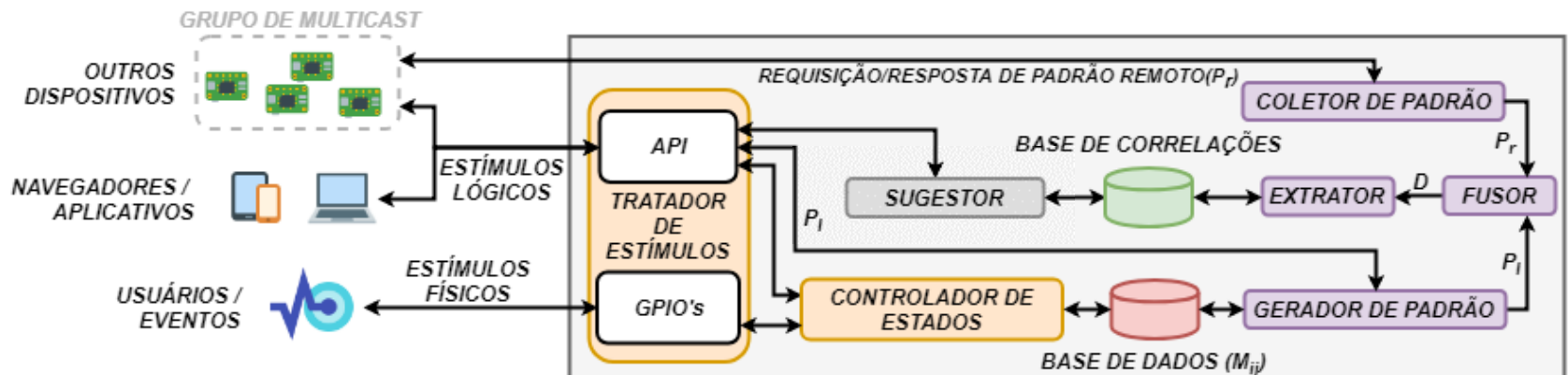


UFAM

Visão geral

I Dispositivos seguem especificações da *Web Thing Model* e compartilham um mesmo grupo de *multicast*

- III • **Extrator:** Realiza a análise associativa na base de transação (D) gerada pelo Fusor e atualiza as regras mais relevantes na base de correlações

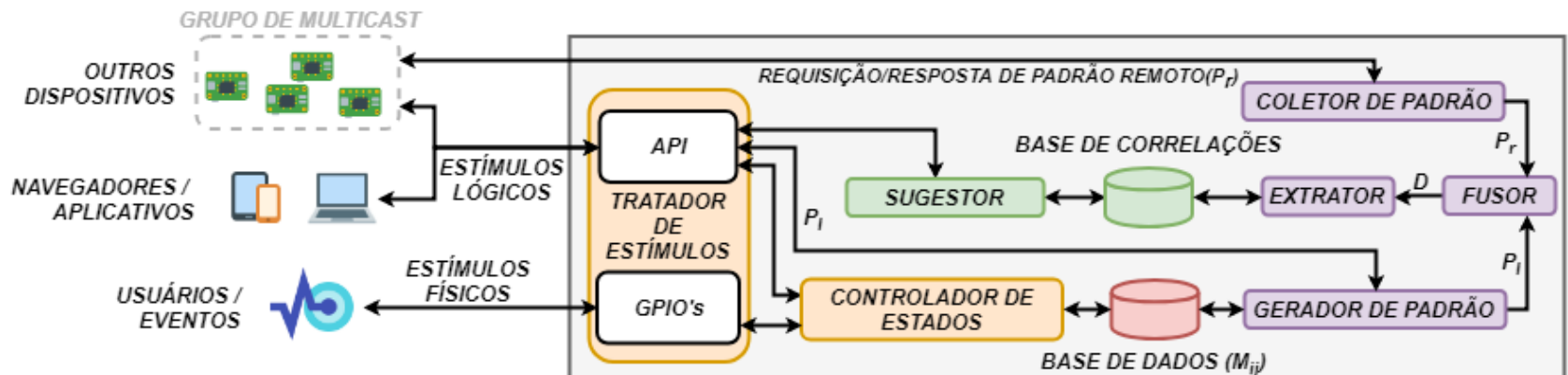


MÉTODO PROPOSTO

Visão geral

I Dispositivos seguem especificações da *Web Thing Model* e compartilham um mesmo grupo de *multicast*

- II
- III
- IV **Sugestor:** Apresenta as regras identificadas ao usuário possibilitando o gerenciamento das mesmas. Indica ao Tratador de Estímulos qual consequente (dispositivo remoto e ação correlata) da regra solicitada
- V
- VI



Métodologia de Avaliação

I

Identificar o quão confiável é o método proposto em relação a análise associativa centralizada

II

- Análise centralizada feita pela *lib* “aRules” no software R
(HAHSLER *et al.* 2011)

III

- Base de transações única (fusão de todos os padrões de ações)

IV

- Identificar se as regras extraídas pelo MAKE também são as mais relevantes extraídas pelo aRules

V

- Regras iguais (*hit*). Regras diferentes (*miss*)

VI

EXPERIMENTOS



Parâmetros dos experimentos

- **Base de dados:** Uma matriz para cada dia da semana;
- **Slots:** Intervalos de 15 minutos ($|T|=96$)
- **Limiares Mínimos:** *Support* (1%), *Lift* (1.1), *Confidence* (90%);

I

II

III

IV

V

VI

EXPERIMENTOS



Parâmetros dos experimentos

- I - Base de dados: Uma matriz para cada dia da semana;
- II - Slots: Intervalos de 15 minutos ($|T|=96$)
- III - Limiares Mínimos: *Support* (1%), *Lift* (1.1), *Confidence* (90%);
- IV - Extrações realizadas para os intervalos:
 - Intervalo I: Diariamente;
 - Intervalo II: Semanas alternadas
 - Intervalo III: Três semanas não e uma semana sim
- V
- VI

EXPERIMENTOS



Parâmetros dos experimentos

- I - **Base de dados:** Uma matriz para cada dia da semana;
- II - **Slots:** Intervalos de 15 minutos ($|T|=96$)
- III - **Limiares Mínimos:** *Support* (1%), *Lift* (1.1), *Confidence* (90%);
- IV - **Extrações realizadas para os intervalos:**
 - **Intervalo I:** Diariamente;
 - **Intervalo II:** Semanas alternadas
 - **Intervalo III:** Três semanas não e uma semana sim
- VI - **Entradas:** WSU-CASAS (COOCK *et al*, 2013):
 - Registro de vários dispositivos(valores contínuos e discretos)
 - Entradas exigiram pré processamento;

EXPERIMENTOS



Exemplo de discretização

I

Baseado nos registros, obtem-se a média dos registros e considera valores maiores ou iguais à média como “HIGH” e abaixo como “LOW”

II

III

IV

V

VI

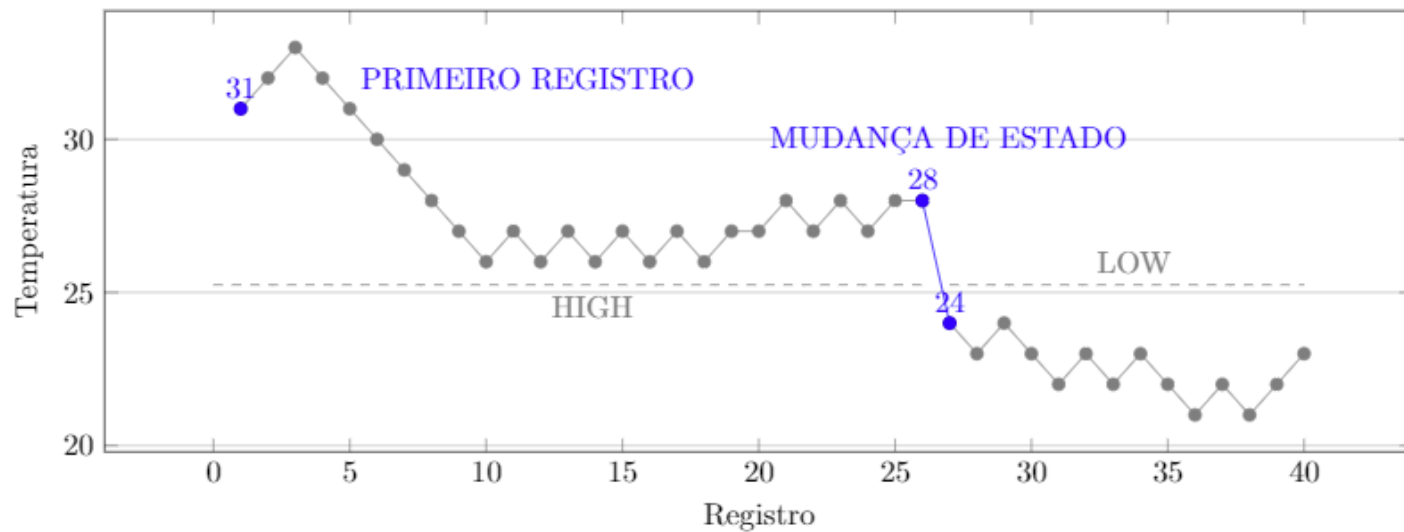
EXPERIMENTOS



UFAM

Exemplo de discretização

I Baseado nos registros, obtem-se a média dos registros e considera valores maiores ou iguais à média como “HIGH” e abaixo como “LOW”



Sensor de temperatura: Média 25.25°C

RESULTADOS



UFAM

Limpeza e discretização

I

Redução massiva no volume de dados que seriam armazenados / processados

II

III

DATASET	DISPOSITIVOS	DIAS	REGISTROS	UTILIZADO	REDUÇÃO
hh107	110	371	3.369.689	2.811.279	16,57%
hh123	88	588	2.907.282	2.345.775	10,31%
hh129	86	668	12.303.984	56.523	99,54%
shib009	8	847	3.187.940	90.599	97,16%
tokyo	67	115	802.534	171.483	78,63%

IV

V

VI

RESULTADOS



UFAM

Limpeza e discretização

Redução massiva no volume de dados que seriam armazenados / processados

DATASET	DISPOSITIVOS	DIAS	REGISTROS	UTILIZADO	REDUÇÃO
hh107	110	371	3.369.689	2.811.279	16,57%
hh123	88	588	2.907.282	2.345.775	10,31%
hh129	86	668	12.303.984	56.523	99,54%
shib009	8	847	3.187.940	90.599	97,16%
tokyo	67	115	802.534	171.483	78,63%

Muitos dispositivos com registros contínuos que “não representavam uma mudança de estado”, considerando a discretização

RESULTADOS



Comparação das Análises

Total regras extraídas durante todo o experimento: **23.073**

DATASET	HITS/MISSES POR INTERVALO			TAXAS MÉDIAS	
	I	II	III	HITS	MISS
hh107	3.656/-	3.493/-	5.769/43	99.67	0.33
hh123	1962/-	2.373/-	4.040/15	99.82	0.18
hh129	80/-	54/-	54/-	100	-
shib009	16/-	135/-	75/-	100	-
tokyo	508/-	337/-	473/-	100	-
TOTAL	23.015 / 58			99.75	0.25

Apenas para duas bases de dados registraram valores miss

Taxa media de hits durante o experimento é de **99.75%**

RESULTADOS



UFAM

Regras não identificadas

Avaliando as regras divergentes, observou-se uma ligeira diferença entre as métricas de *support* e *lift* em ambas análises

Antecedente	Consequente	MAKE (Misses)			<i>aRules</i> (Permissivo)			Contagem
		Supp	Conf	Lift	Supp	Conf	Lift	
LS023 (HIGH)	LS021(LOW)	0.5729	0.9821	1.1359	0.5978	0.9821	<u>1.0886</u>	55
LS019 (LOW)	LS021(LOW)	0.7395	0.9861	1.1405	0.7717	0.9861	<u>1.0930</u>	71
LS004 (HIGH)	LS021(LOW)	0.6250	0.9836	1.1376	0.6521	0.9836	<u>1.0902</u>	60

RESULTADOS



Regras não identificadas

Avaliando as regras divergentes, observou-se uma ligeira diferença entre as métricas de *support* e *lift* em ambas análises

Antecedente	Consequente	MAKE (Misses)			<i>aRules</i> (Permissivo)			Contagem
		Supp	Conf	Lift	Supp	Conf	Lift	
LS023 (HIGH)	LS021(LOW)	0.5729	0.9821	1.1359	0.5978	0.9821	<u>1.0886</u>	55
LS019 (LOW)	LS021(LOW)	0.7395	0.9861	1.1405	0.7717	0.9861	<u>1.0930</u>	71
LS004 (HIGH)	LS021(LOW)	0.6250	0.9836	1.1376	0.6521	0.9836	<u>1.0902</u>	60

Diferença gerada pela adaptação do cálculo das métricas $|D| = |T|$
aRules com $|D| < |T|$ que interfere diretamente em *lift* e *supp*

RESULTADOS



UFAM

Regras não identificadas

Avaliando as regras divergentes, observou-se uma ligeira diferença entre as métricas de *support* e *lift* em ambas análises

Antecedente	Consequente	MAKE (Misses)			<i>aRules</i> (Permissivo)			Contagem
		Supp	Conf	Lift	Supp	Conf	Lift	
LS023 (HIGH)	LS021(LOW)	0.5729	0.9821	1.1359	0.5978	0.9821	<u>1.0886</u>	55
LS019 (LOW)	LS021(LOW)	0.7395	0.9861	1.1405	0.7717	0.9861	<u>1.0930</u>	71
LS004 (HIGH)	LS021(LOW)	0.6250	0.9836	1.1376	0.6521	0.9836	<u>1.0902</u>	60

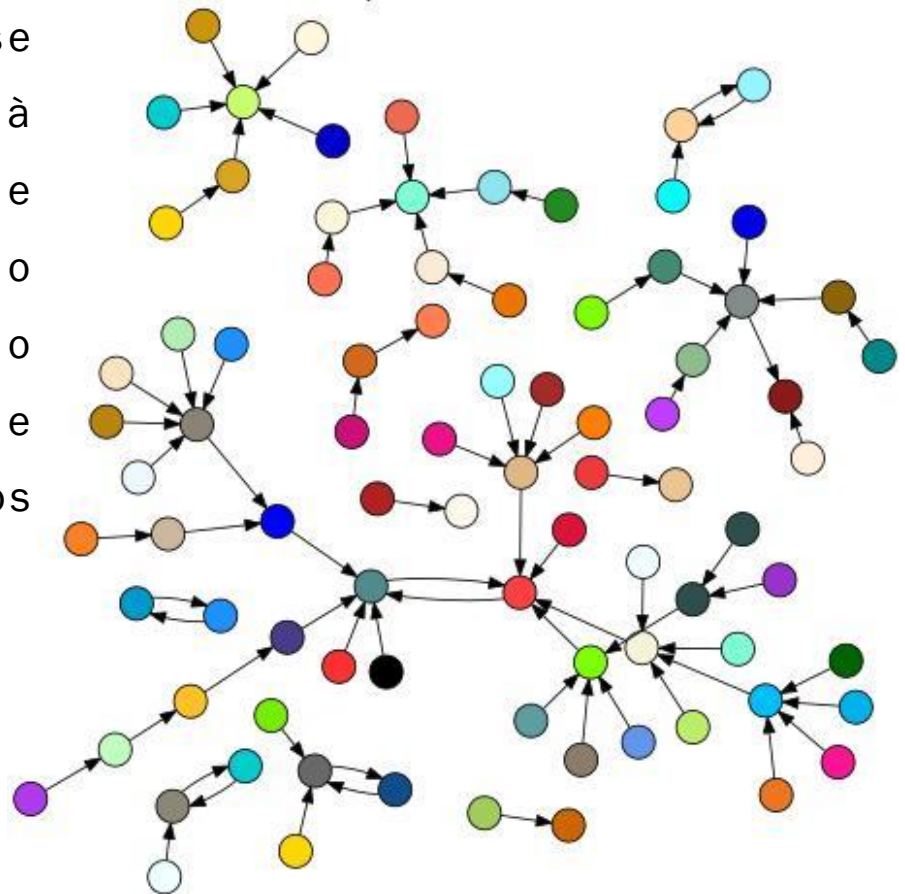
Diferença gerada pela adaptação do cálculo das métricas $|D| = |T|$ aRules com $|D| < |T|$ que interfere diretamente em *lift* e *supp*

Regras presentes em mais da metade dos *slots*. Apenas o MAKE identificou como relevante

CONCLUSÕES

Mineração descentralizada e embarcada.

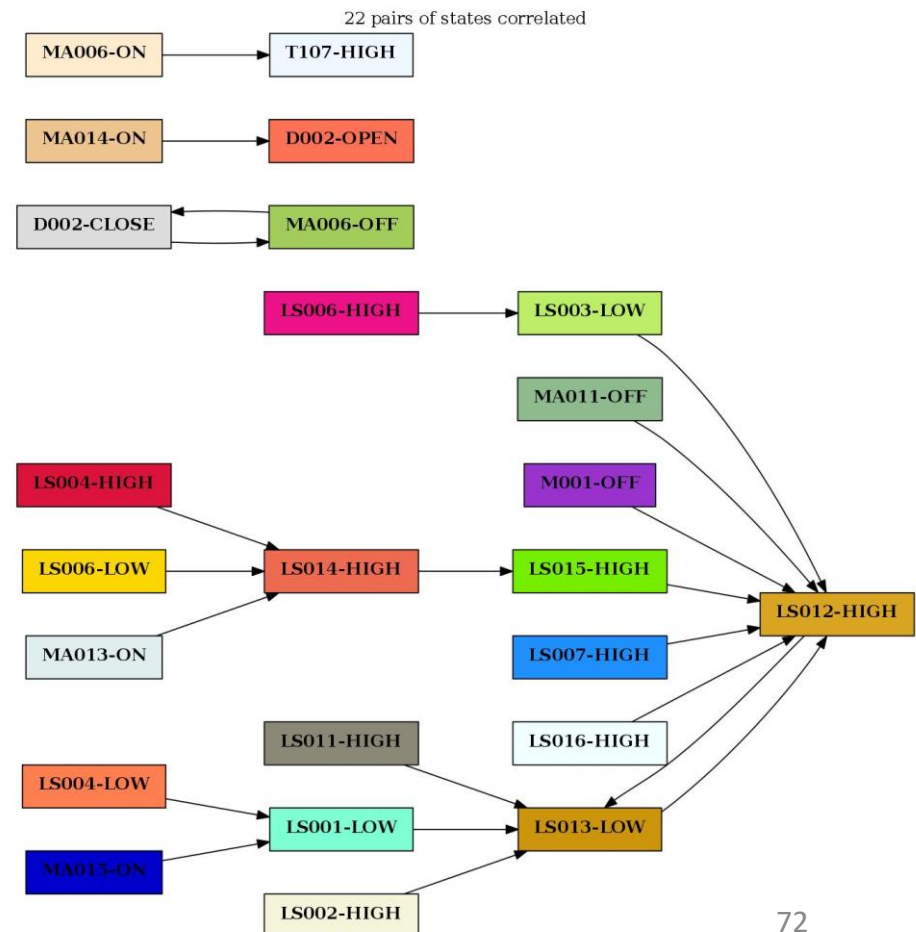
- I - O método proposto mostrou-se
- II eficiente quanto à
- III identificação de regras de
- IV correlação em cenário
- V distribuído, dispensando o
- VI uso de servidores e
- dispositivos com muitos
- recursos



CONCLUSÕES

Mineração descentralizada e embarcada.

- I - O registro e análise de
 - II dados baseada nas ações do
 - III dispositivo apresentou
 - IV resultados interessantes,
 - V tanto para a redução de
 - VI dados quanto para a
- capacidade de armazenar informações relevantes entre os dispositivos



CONCLUSÕES



UFAM

Mineração descentralizada e embarcada.

- O armazenamento de dados em forma de matriz de contadores demonstrou ser uma alternativa viável para dispositivos com poucos recursos podendo ser facilmente dimensionável para se ajustar às limitações do dispositivo

Diagram illustrating the data storage and processing flow:

Base de dados (Matriz) (Database Matrix) is connected to **Padrão de ações** (Action Pattern).

The **Padrão de ações** is shown in a JSON format, with tabs for **JSON**, **Raw Data**, and **Headers**. The **Raw Data** tab is selected, showing the following JSON structure:

```
{
  "actions": ["off", "on"],
  "dataset": {
    "42": [3.6, 3.6],
    "43": [0, 1.6],
    "41": [2.83, 2.6],
    "5": [2.6, 1],
    "46": [2.83, 3],
    "45": [1.6, 0],
    "35": [2, 3],
    "48": [0, 3],
    "16": [1, 3.33],
    "23": [3.17, 1],
    "12": [1, 1.6],
    "10": [1.6, 3.17],
    "3": [3, 2.6],
    "2": [1.6, 1],
    "19": [2.83, 1],
    "6": [1.6, 1.6],
    "14": [2, 2.83],
    "24": [1, 2],
    "27": [2, 1],
    "26": [1, 2.83],
    "30": [2, 1.6],
    "33": [0, 2.6],
    "32": [2.33, 3.33]
  }
}
```

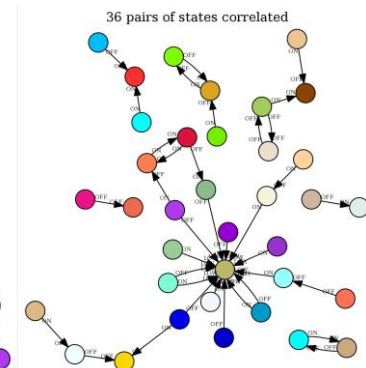
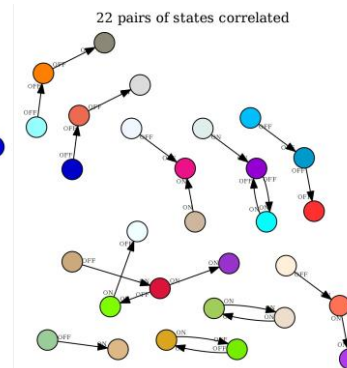
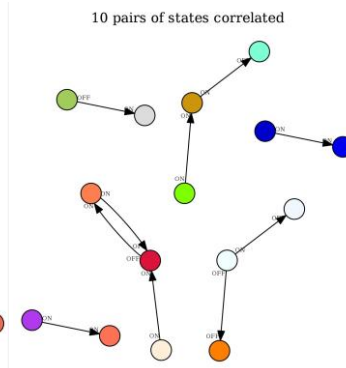
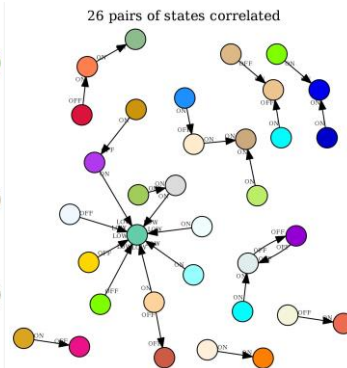
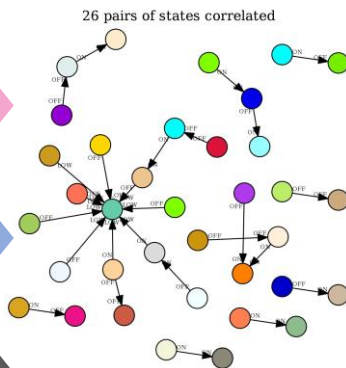
The **Base de dados (Matriz)** is shown in a JSON format, with tabs for **JSON**, **Raw Data**, and **Headers**. The **Raw Data** tab is selected, showing the following JSON structure:

```
{
  "actions": ["off", "on"],
  "pattern": {
    "27": 0,
    "3": 0,
    "10": 1,
    "5": 0,
    "46": 1,
    "19": 0,
    "35": 1,
    "48": 1,
    "24": 1,
    "16": 1,
    "26": 1,
    "32": 1,
    "30": 0,
    "33": 1,
    "23": 0
  }
}
```

CONCLUSÕES

Mineração descentralizada e embarcada.

- As extrações de intervalos de tempos fixos possibilitaram a identificação no padrão de ações de tal forma que foi possível adaptar as correlações ao longo do tempo



TOKYO – INTERVALO III – QUARTA – CHECKPOINT 01 à 05

CONCLUSÕES



Observações Importantes

- I • Dimensionalidade da Matriz ($|A| \times |T|$): A dimensão da matriz deve ser avaliada com cautela já que se o número de *slots* e/ou o número de ações do dispositivo for elevado, a mesma pode consumir muito espaço (de armazenamento e processamento)
- II
- III
- IV
- V
- VI

CONCLUSÕES



Observações Importantes

- I** • **Dimensionalidade da Matriz ($|A| \times |T|$):** A dimensão da matriz deve ser avaliada com cautela já que se o número de *slots* e/ou o número de ações do dispositivo for elevado, a mesma pode consumir muito espaço (de armazenamento e processamento).
- II**
- III**
- IV** • **Compartilhamento de padrões:** Todos os dispositivos devem ser capazes conciliar dois padrões de ações simultaneamente para que seja possível executar a análise associativa. Para matrizes com muitos *slots*, isso pode ser um problema
- V**
- VI**

CONCLUSÕES



Observações Importantes

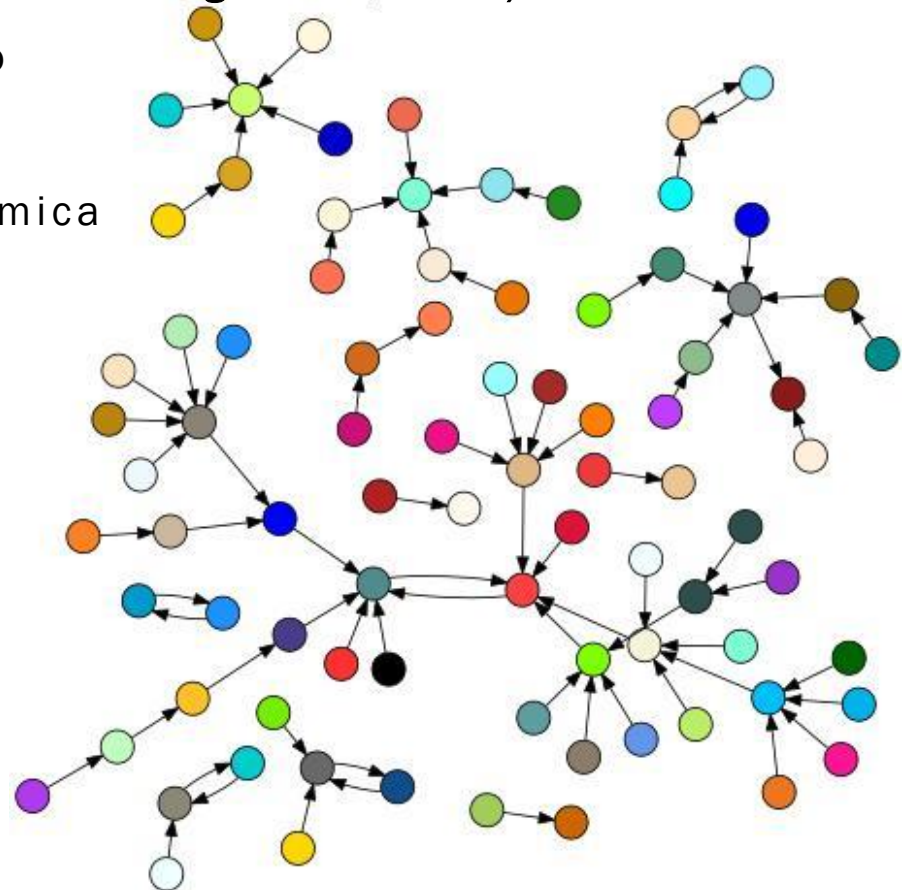
- I** • **Dimensionalidade da Matriz ($|A| \times |T|$):** A dimensão da matriz deve ser avaliada com cautela já que se o número de *slots* e/ou o número de ações do dispositivo for elevado, a mesma pode consumir muito espaço (de armazenamento e processamento).
- II**
- III**
- IV** • **Compartilhamento de padrões:** Todos os dispositivos devem ser capazes conciliar dois padrões de ações simultaneamente para que seja possível executar a análise associativa. Para matrizes com muitos *slots*, isso pode ser um problema.
- V**
- VI**
- **Descoberta de dispositivos:** Protocolo multicast usa UDP

CONCLUSÕES

Considerações Finais

- MAKE (eMbedded Associative Knowledge Extraction)

- ✓ Ambiente descentralizado
- ✓ Dispositivos autônomos
- ✓ Mineração de dados dinâmica
- ✓ Modelo adaptativo
- ✓ Fácil Integração (WTM)





Universidade Federal do Amazonas
Instituto de Computação
Programa de Pós-Graduação em Informática
Grupo de Interesse em Sistemas Embarcados

DÚVIDAS

Manaus, 25 de março de 2019



Universidade Federal do Amazonas
Instituto de Computação
Programa de Pós-Graduação em Informática
Grupo de Interesse em Sistemas Embarcados

OBRIGADO

Manaus, 25 de março de 2019