

# **ADMin: Adaptive monitoring dissemination for the Internet of Things**

Demetris Trihinas, George Pallis, Marios D. Dikaiakos

IEEE International Conference on Computer Communications - INFOCOM

2017

# Sumário

- Introdução
- Proposta ADMin
- Resultados
- Pontos fortes e fracos

- Motivação:
  - Dispositivos de IoT geralmente possuem restrição de energia e de taxa para transmissão ou recepção de dados;
  - ADMin tem o objetivo de reduzir a quantidade de dados transmitidos pelos dispositivos IoT e conseqüentemente o consumo de energia;

- Características gerais do ADMin:
  - Reduz a quantidade de dados transmitidos com o uso de técnicas de estimativas dos dados coletados pelos sensores;
  - Comparado com outras propostas, possui o diferencial de introduzir dados de tendência e sazonalidade para aumentar a precisão das estimativas;

- Características gerais do ADMin:
  - Reduz a quantidade de dados transmitidos com o uso de técnicas de estimativas dos dados coletados pelos sensores;
  - Comparado com outras propostas, possui o diferencial de introduzir dados de tendência e sazonalidade para aumentar a precisão das estimativas;

Open Source

- Características gerais do ADMin:
  - Reduz a quantidade de dados transmitidos com o uso de técnicas de estimativas dos dados coletados pelos sensores;
  - Comparado com outras propostas, possui o diferencial de introduzir dados de tendência e sazonalidade para aumentar a precisão das estimativas;

Open Source

Tempo Real

- Fluxos de medições é composto por uma sequência de pontos de dados  $M$ :

$$M = \{d_i\}_{i=0}^n \quad n \rightarrow \infty$$

- Pontos de dados medidos são representados por tuplas:

$$d_i = (e_{id}, t_i, v_i, \dots)$$

onde:

$e_{id}$  = identificador único do dado;

$t_i$  = *timestamp*

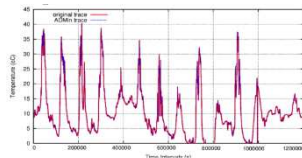
$v_i$  = valor medido

# Introdução

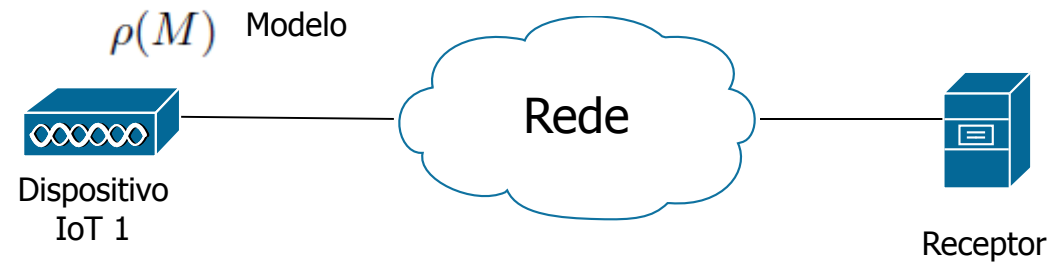
## Arquitetura do ADMIN

$$M = \{d_i\}_{i=0}^n \quad n \rightarrow \infty$$

Fluxo de dados



Grandeza medida  
periodicamente a cada T  
unidades de tempo



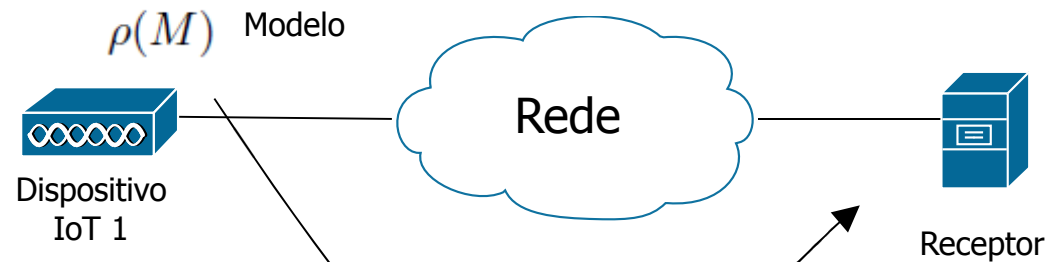
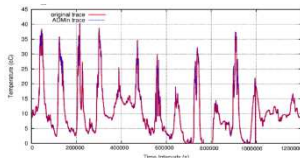


# Introdução

## Arquitetura

$$M = \{d_i\}_{i=0}^n \quad n \rightarrow \infty$$

Fluxo de dados



$$d_{i+k|i} = f(\rho(M), d_i)$$

Instante  $t_i$  =  
disseminação  
inicial

$$t_k = t_i + k \cdot T \quad | \quad k \subseteq \mathbb{Z}^+$$

# Introdução

- Precisão do modelo de estimativa usado em ADMin:

- Precisão do modelo de estimativa usado em ADMin:

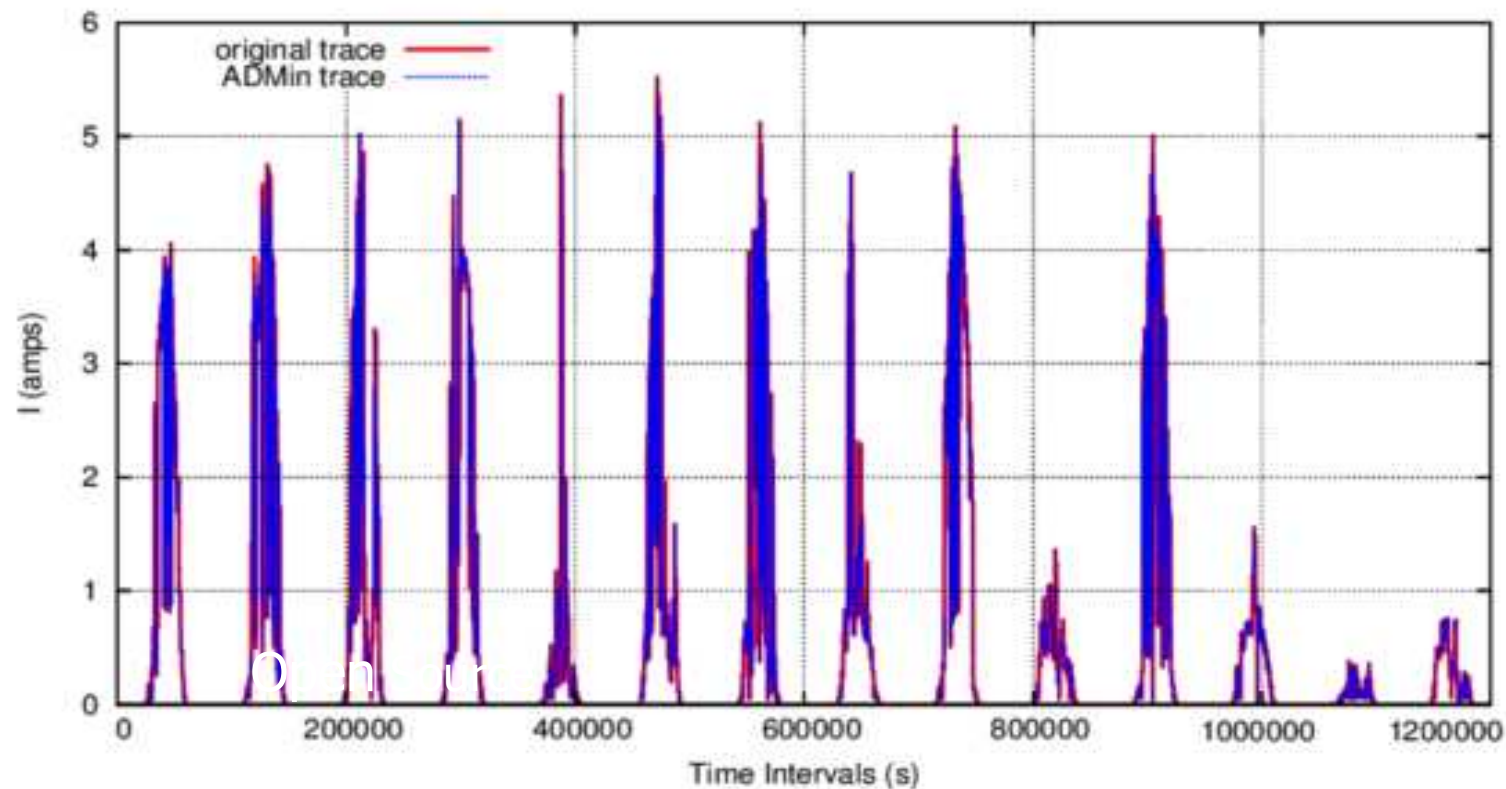


Fig. 1: PV Panel Current Production Trace

- Precisão do modelo de estimativa usado em ADMin:

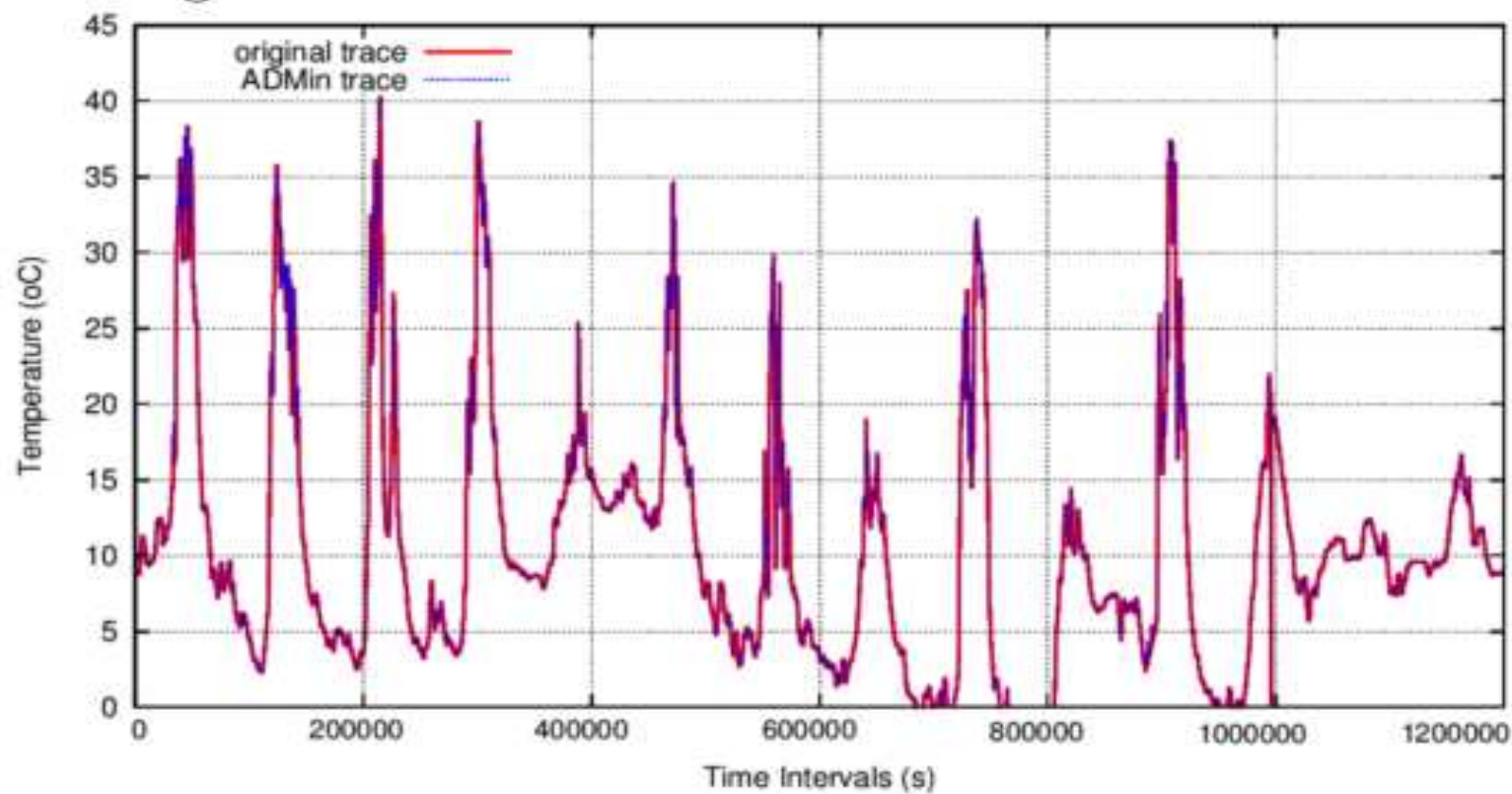
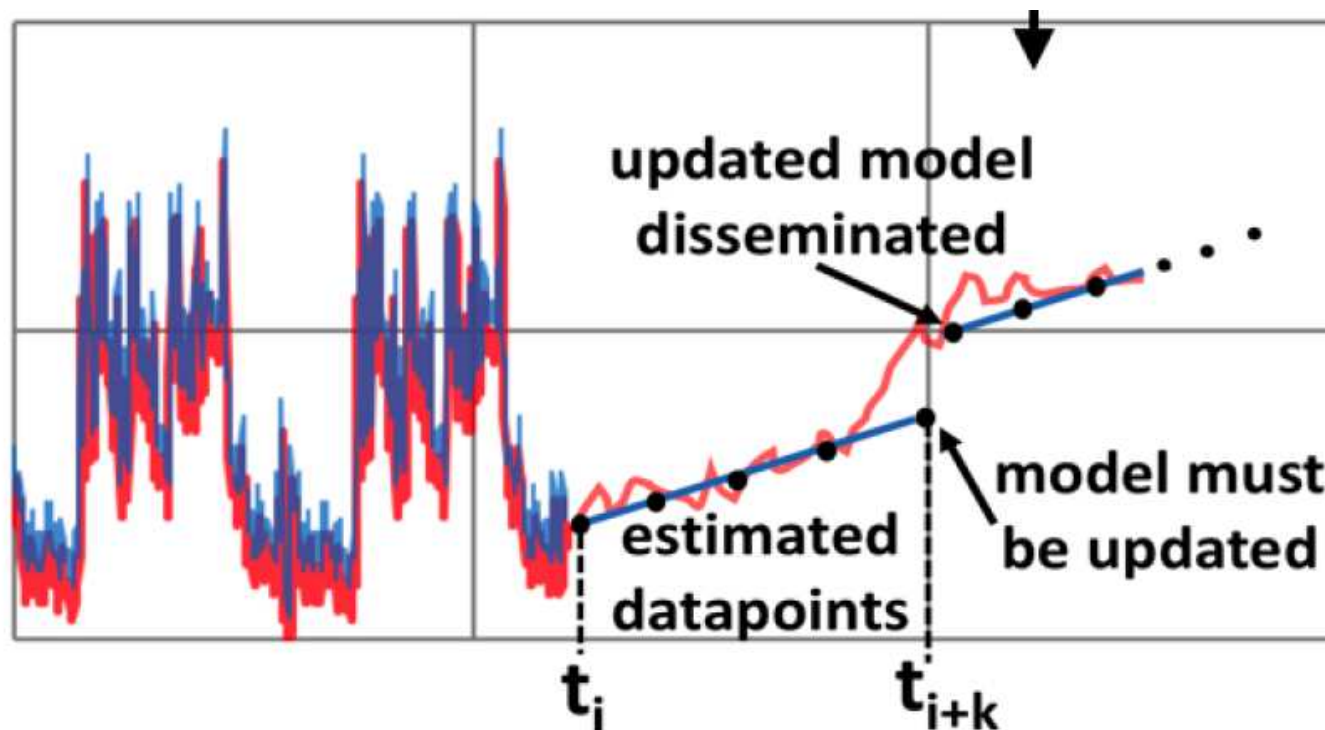


Fig. 2: Weather Station Temperature Trace

# Proposta ADMin

## Visão geral da monitoração adaptativa:

Suprime o envio de pontos de dados com “pequenas” variações com relação aos pontos prévios. Modelo de estimativa é enviado no início e quando há decisão na fonte para atualização do modelo;



# Proposta ADMin

---

Visão geral da monitoração adaptativa:

Requerimentos:

- **R1:** O processo de estimativa deve ser processado em código leve e executado no local da fonte de monitoração;

Visão geral da monitoração adaptativa:

Requerimentos:

- **R1:** O processo de estimativa deve ser processado em código leve e executado no local da fonte de monitoração;
- **R2:** O processo de estimativa deve ser eficiente. Gastar menos energia do que o que seria transmitir os pontos de dados e descartá-los

Visão geral da monitoração adaptativa:

Requerimentos:

- **R1:** O processo de estimativa deve ser processado em código leve e executado no local da fonte de monitoração;
- **R2:** O processo de estimativa deve ser eficiente. Gostar menos energia do que o que seria transmitir os pontos de dados e descartá-los
- **R3:** Enquanto os parâmetros da arquitetura puderem ser ajustados automaticamente, nenhum usuário deve entrar “parâmetros mágicos” no sistema para o correto funcionamento;



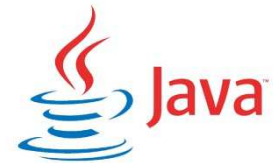
Visão geral da monitoração adaptativa:

Requerimentos:

- **R1:** O processo de estimativa deve ser processado em código leve e executado no local da fonte de monitoração;
- **R2:** O processo de estimativa deve ser eficiente. Gostar menos energia do que o que seria transmitir os pontos de dados e descartá-los
- **R3:** Enquanto os parâmetros da arquitetura puderem ser ajustados automaticamente, nenhum usuário deve entrar “parâmetros mágicos” no sistema para o correto funcionamento;
- **R4:** O sistema ADMin deve ser prático, atingindo bom desempenho em diversos *Testbeds* diferentes;

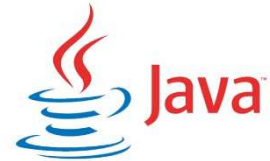
# Proposta ADMin

Desenvolvido em Java



# Proposta ADMin

Desenvolvido em Java



Dispositivos com flexibilidade  
para rodar o código



Ex: Raspberry pi

# Proposta ADMin

Desenvolvido em Java



Dispositivos com flexibilidade  
para rodar o código

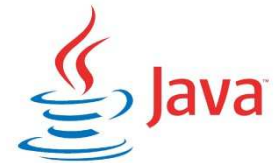


Ex: Raspberry pi

Arquitetura Modular

# Proposta ADMin

Desenvolvido em Java



Dispositivos com flexibilidade para rodar o código



Ex: Raspberry pi

Arquitetura Modular

Adaptative Stream Estimation

# Proposta ADMin

Desenvolvido em Java



Dispositivos com flexibilidade para rodar o código



Ex: Raspberry pi

Arquitetura Modular

Adaptative Stream Estimation

Seasonality Enrichment

# Proposta ADMin

Desenvolvido em Java



Dispositivos com flexibilidade para rodar o código



Ex: Raspberry pi

Arquitetura Modular

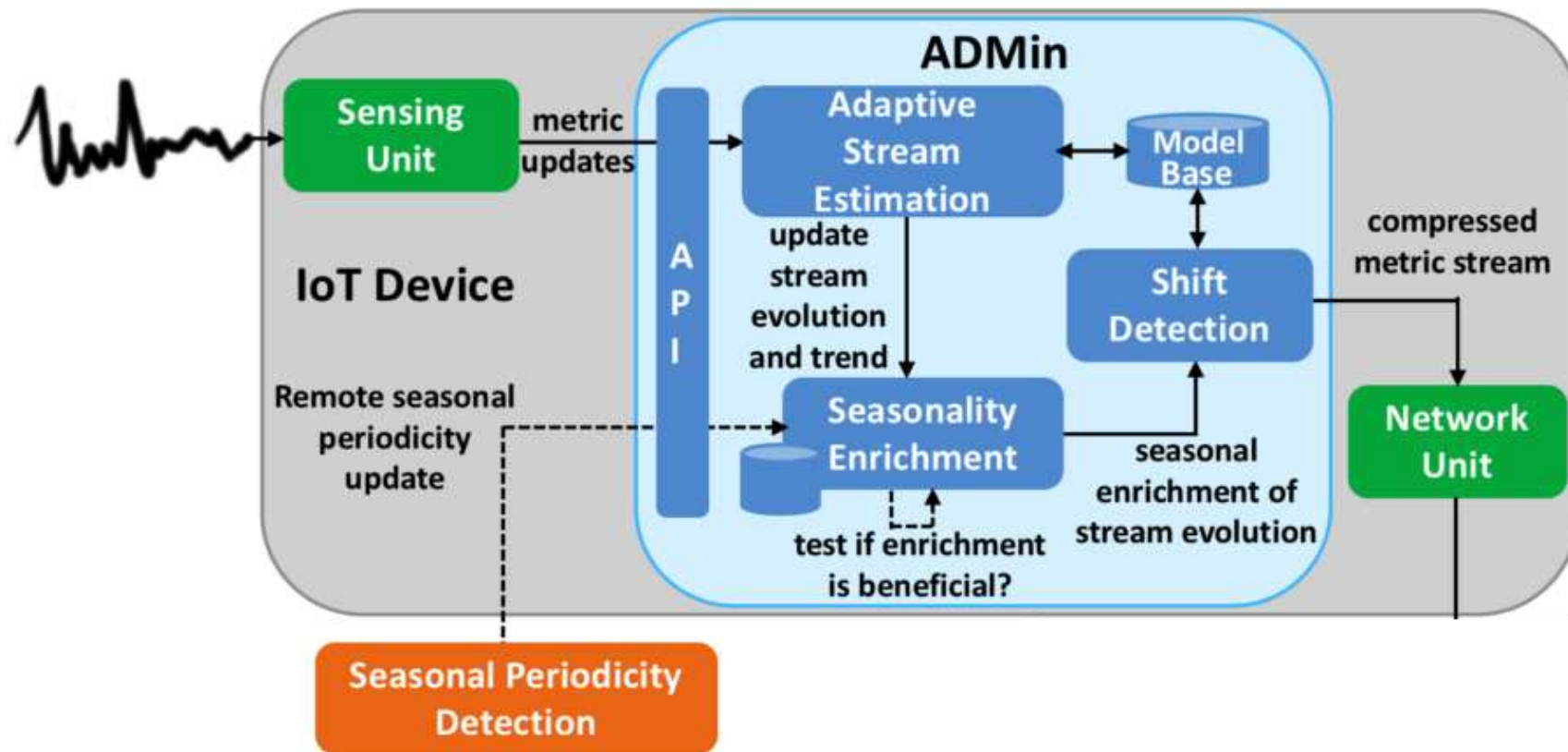
Adaptative Stream Estimation

Seasonality Enrichment

Shift Detection

# Proposta ADMin

Arquitetura Modular ADMIN:





## Arquitetura Modular ADMin: Adaptive Stream Estimation

- Classifica os pontos de dados estimados como “esperados” ou “não esperados”;
- Pontos de dados “esperados” são os que podem ser determinados pelo modelo de estimativa. Estes pontos são descartados;
- Pontos de dados “não esperados” são armazenados para posterior envio, quando a disseminação dos dados é disparada;

## Arquitetura Modular ADMin: Adaptive Stream Estimation

- Método utilizado como base: *Exponential Weighted Moving Average* (EWMA)

$$\mu_i = \alpha\mu_{i-1} + (1 - \alpha)v_i \quad (1)$$

$\alpha$  = Fator de introduzido para reduzir exponencialmente a influência de valores antigos

$\mu_i$  = Média móvel

## Arquitetura Modular ADMin: Adaptive Stream Estimation

- A proposta ADMin utiliza uma evolução da EWMA, chamada PEWMA (*Probabilistic Exponential Weighted Moving Average*)

$$\mu_i = \begin{cases} v_i, & i = 1 \\ \alpha(1 - \beta P_i)\mu_{i-1} + (1 - \alpha(1 - \beta P_i))v_i, & i > 1 \end{cases} \quad (2)$$

$\beta$ . = Fator para introduzir a tendência na estimativa, melhorando a precisão na identificação de mudanças bruscas

## Arquitetura Modular ADMin: Adaptive Stream Estimation

- A proposta ADMin utiliza uma evolução da EWMA, chamada PEWMA (*Probabilistic Exponential Weighted Moving Average*)

$$P_i = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z_i^2}{2}\right) \quad (3)$$
$$Z_i = \frac{v_i - \hat{v}_i}{\sigma_i}$$

$P_i$  é a probabilidade de  $v_i$ , considerando uma distribuição Gaussiana.

## Arquitetura Modular ADMin: Adaptive Stream Estimation

- A proposta ADMin utiliza uma evolução da EWMA, chamada PEWMA

( $P_i$ )

$(1 - \beta P_i)$  Permite que valores inesperados e repentinos sejam considerados no processo de estimativa mas oferecendo baixo peso para esses cálculos;

No entanto, quando há uma mudança brusca real na tendência, este fator permite que os novos valores medidos aumentem a probabilidade de acerto dos próximos valores estimados;

precisão na identificação de mudanças bruscas

## Arquitetura Modular ADMin: Adptative Stream Estimation

- Somente usar a média móvel não é suficiente para estimativas precisas. A proposta acrescenta, portanto, a **tendência** ( $X_i$ ) nos cálculos:

## Arquitetura Modular ADMin: Adaptive Stream Estimation

- Somente usar a média móvel não é suficiente para estimativas precisas. A proposta acrescenta, portanto, a **tendência** ( $X_i$ ) nos cálculos:

$$X_i = \begin{cases} v_i - v_{i-1}, & i = 2 \\ \gamma (\mu_i - \mu_{i-1}) + (1 - \gamma) X_{i-1}, & i > 2 \end{cases} \quad (4)$$

$\gamma$   $\rightarrow$  Parâmetro utilizado para amortizar a tendência. Varia entre 0 e 1.

## Arquitetura Modular ADMin: Adaptive Stream Estimation

- O principal objetivo no acréscimo da tendência é diminuir o efeito de atrasos no processo de estimativa;
- Assim, os valores dos pontos de dados estimados podem ser calculados pela fórmula:

$$\hat{v}_{i+k|i} = \mu_i + k X_i \quad (5)$$



## Arquitetura Modular ADMin: Seasonality Enrichment

- Detecta se é vantajoso ou não atualizar o modelo com informações de Sazonalidade;
- Sazonalidade é definida como a tendência do fluxo de medições exibir um comportamento que se repete a cada  $L$  períodos (ex: diariamente, semanalmente);

## Arquitetura Modular ADMin: Seasonality Enrichment

- Para cálculo do fator sazonalidade a proposta usa o método de Holt-Winter

$$S_i = \begin{cases} 0, & i < L \\ \omega (v_i - \mu_i - X_i) + (1 - \omega) (v_i - S_{i-L}), & i > L \end{cases} \quad (6)$$

$S_{i-L}$  → Fator sazonal do último período;

$\omega$  → Fator de amortecimento. Varia entre 0 e 1.

## Arquitetura Modular ADMin: Seasonality Enrichment

- Assim, os valores dos pontos de dados estimados passam a ser calculados pela fórmula:

$$\hat{v}_{i+k|i} = \mu_i + k X_i + S_i \quad (7)$$

## Arquitetura Modular ADMin: Seasonality Enrichment

- Assim, os valores dos pontos de dados passam a ser calculados pela fórmula:

$$\hat{v}_{i+k|i} = \mu_i + k X_i + S_i \quad (7)$$

- A proposta utiliza dois testes estatísticos (t-tests) para avaliar se o fator sazonalidade irá permitir estimativas mais precisas ou não;

## Arquitetura Modular ADMin: Adaptive Shift Detection

- Utiliza o método *Cumulative Sum Test* (CUSUM),  $C_i$ , que é um teste de hipótese para detectar deslocamentos em séries temporais;

$$c_i = \ln \frac{P(M_i, \theta'')}{P(M_i, \theta')} \quad (8)$$

$$C_{i, \{low, high\}} = C_{i-1, \{low, high\}} + c_i$$

$P(M, \theta')$  → Distribuição estatística das medidas antes do deslocamento, ts

$P(M, \theta'')$  → Distribuição estatística das medidas após do deslocamento, ts

*low*, *high* indicam se os deslocamentos são crescentes ou decrescentes

## Arquitetura Modular ADMin: Adaptive Shift Detection

- O momento em que o deslocamento ocorre é determinado por:

$$G_{i,,\{low, high\}} = \{G_{i-1,\{low, high\}} + c_i\}^+$$
$$t_s = \arg \min_{j \leq s \leq i} (C_{s-1}) \quad (10)$$

$$z^+ = \sup(z, 0)$$

$t_i$  é momento quando ADMIN detecta a mudança e  $t_j$  é o momento quando é registrado o último ponto de dado antes da mudança em  $t_s$ .

## Arquitetura Modular ADMin: Adaptive Shift Detection

- O momento em que o deslocamento ocorre é determinado por:

$$G_{i,,\{low, high\}} = \{G_{i-1,\{low, high\}} + c_i\}^+$$
$$t_s = \arg \min_{j \leq s \leq i} (C_{s-1}) \quad (10)$$

$$z^+ = \sup(z, 0)$$

Se  $G_i$  for maior que um limiar ( $h$ ), o deslocamento é registrado e a disseminação dos pontos de dados "não esperados" é feito. Nesse momento, o modelo também é atualizado com um novo valor de  $h$ .

$t_i$  é momento quando ADMIN detecta a mudança e  $t_j$  é o momento quando é registrado o último ponto de dado antes da mudança em  $t_s$ .

Compara ADMin com outras arquiteturas estado-da-arte para IoT: G-SIP, LANCE e ADWin.

- G-SIP usa EWMA como modelo de estimativa. Disseminação depende da taxa de variação no tempo do fluxo de medições;
- LANCE também usa EWMA. A disseminação é feita quando o número de pontos estimados ultrapassa um threshold;
- ADWin usa modelo de estimativa baseado em Bayes-simples (*Naive Bayes*). Disseminação a mesma técnica de threshold com janela de observação do LANCE;



# Resultados

- Parâmetros:
  - Escolheu o intervalo de confiança  $\delta$  em 0,9;
  - Para as soluções que usam média móvel:  $\alpha = 0.45$ . Porque é a melhor configuração para o G-SIP e para o LANCE;
  - Para Tendência e Sazonalidade (ADMin e G\_SIP):  $\gamma = 0.95$  e  $\omega = 0.35$  (valores default)

# Resultados

- Hardware:
  - Para “PV” e “Temperatura” usou um Raspberry (1ª geração) com 512 MB de RAM e processador ARM (1 core de 700 MHz);
  - O teste “Heartrate” utilizou Android Wear Emulator com um app BPMs para registrar batimentos (processador ARM 23 MHz, 128 MB de RAM);

# Resultados

- Dados:

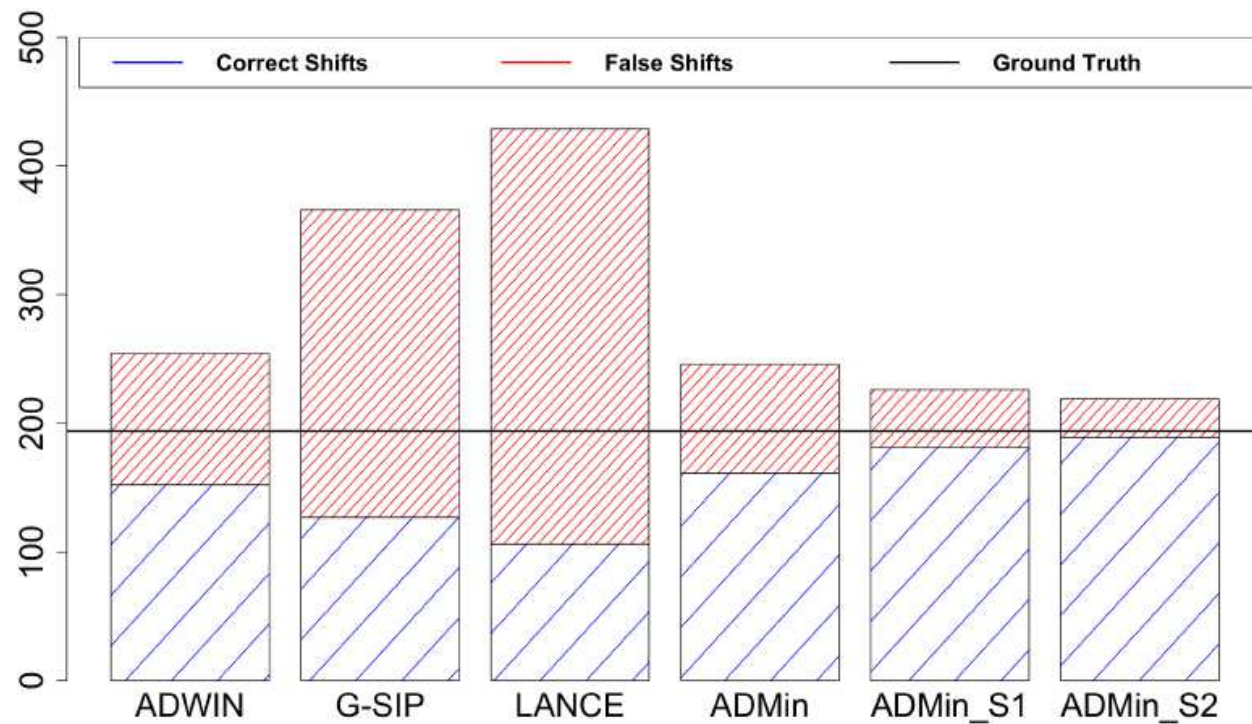
Trace Name	Origin	Data Points	Optimal Shifts	Description
PV Current	PV Panel	1209598	194	A Photovoltaic (PV) current production trace collected from a PV panel every 1 second for a period of 2 weeks in Jan 2015
Temperature	Meteo Station	1209598	572	A Temperature trace collected from a remote weather station monitoring the temperature every 1 second for a period of 2 weeks in Jan 2015
Heartrate	Wearable	40908	202	A Heartrate trace collected from a Fitbit HR wearable device monitoring beats per minute (bpm) of the person wearing the device for a month (Jun 2016)

TABLE I: Traces Used for Performance and Accuracy Evaluation

- Configurações:
  - ADMin significa a configuração sem considerar Sazonalidade;
  - ADMin\_S1 significa a configuração com Sazonalidade estática, definida apenas na inicialização;
  - ADMin\_S2 significa a configuração com Sazonalidade atualizada em tempo real, informada pelo sistema externo ComCube;

# Resultados

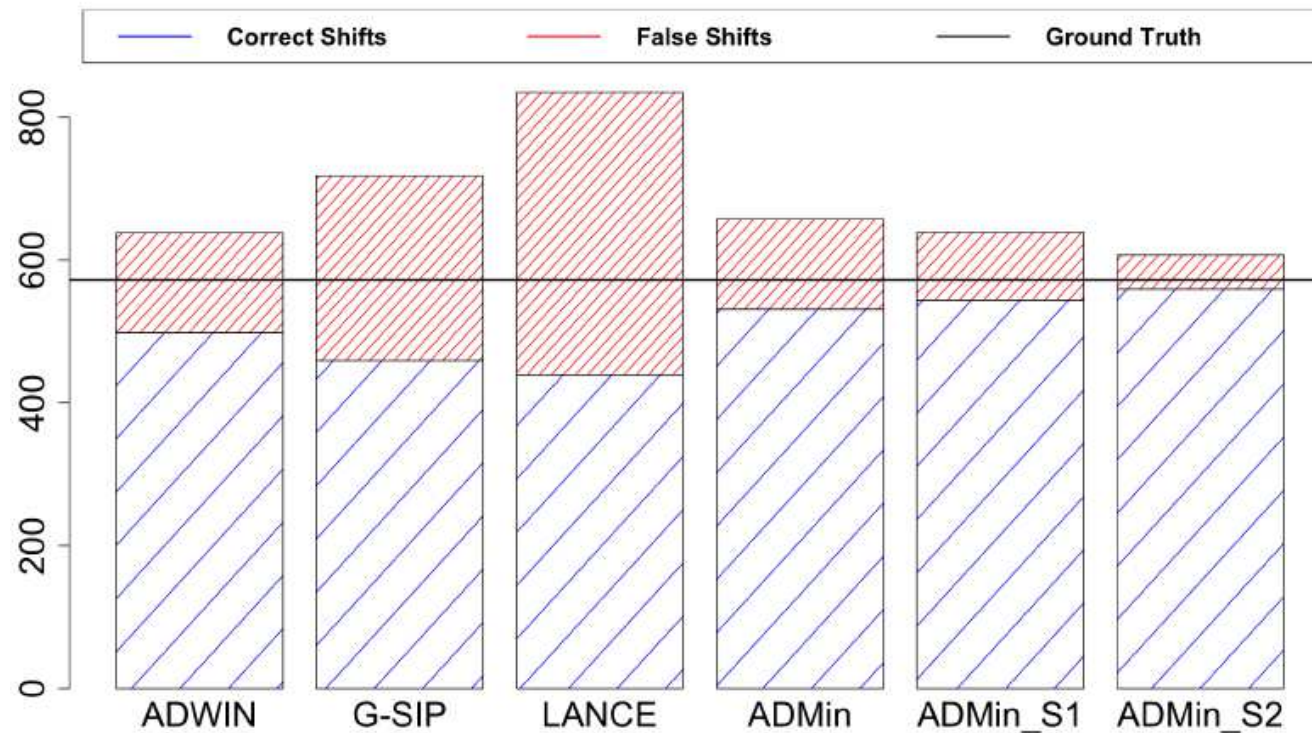
- Precisão:



(a) PV Current Trace

# Resultados

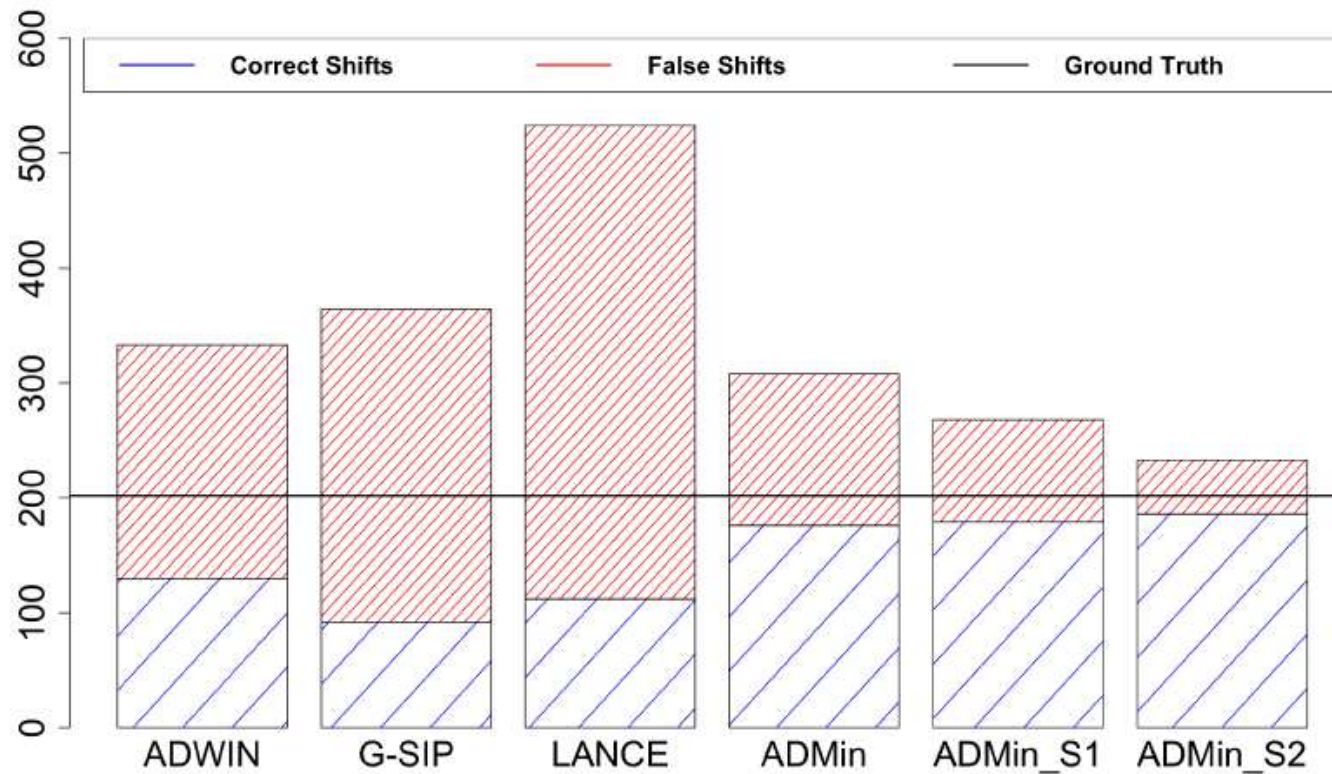
- Precisão:



(b) Weather Station Temperature Trace

# Resultados

- Precisão:



(c) Heartrate Trace

# Resultados

- Consumo de energia:

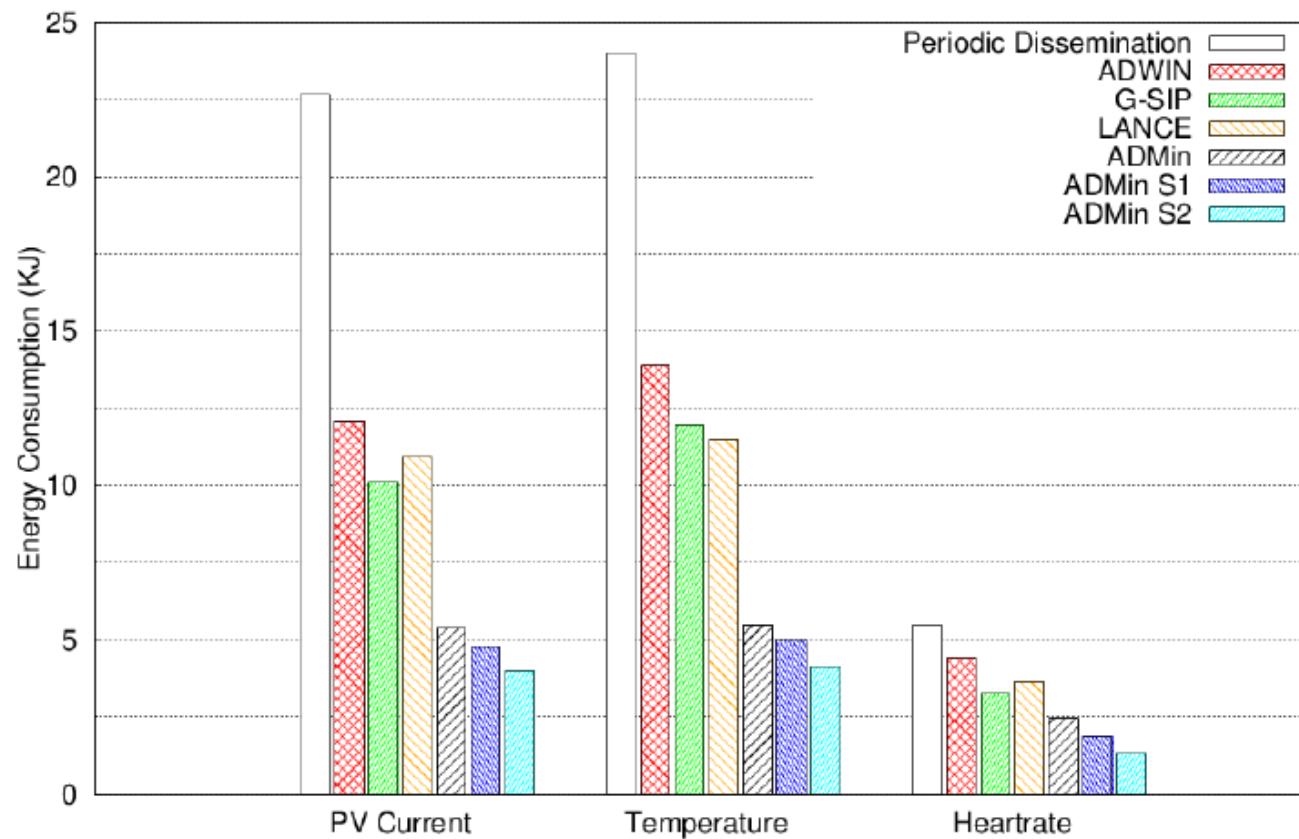


Fig. 6: On Device Energy Consumption Comparison



# Resultados

- Redução no volume de dados enviados:

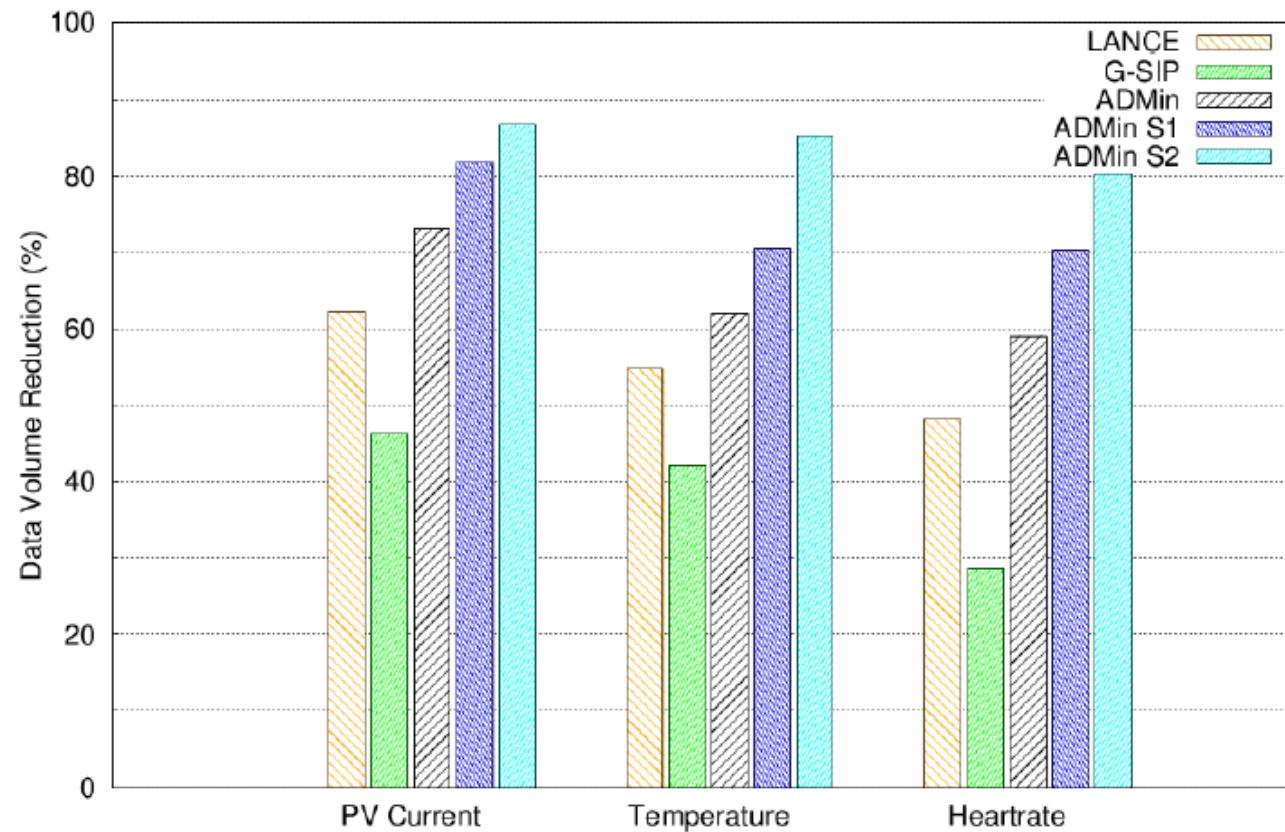


Fig. 7: Data Reduction Comparison

# Resultados

- Precisão da arquitetura do ponto de vista do receptor

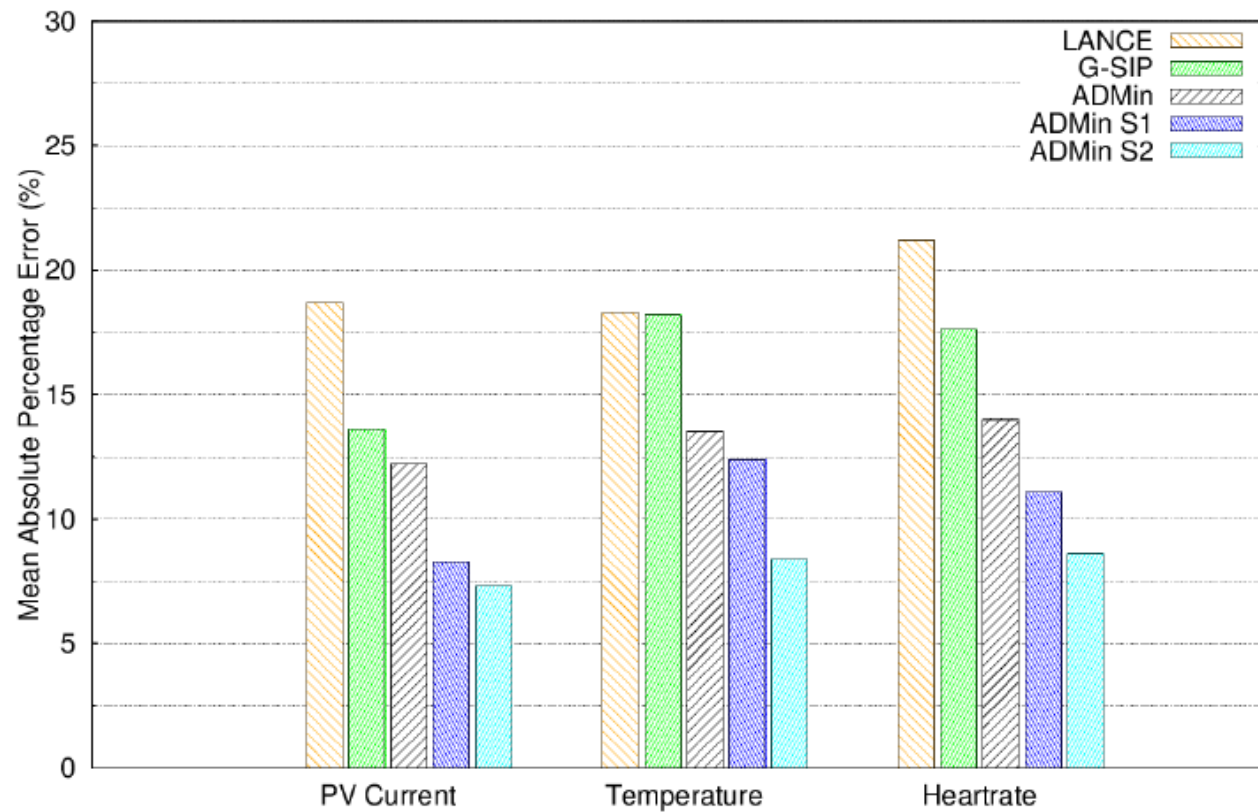


Fig. 8: Receiver-Side Mean Absolute Percentage Error

# Pontos fortes e fracos

- Pontos fortes e fracos do artigo:

Pontos fortes	Pontos fracos
Apresenta solução para um problema relevante dentro do campo específico	Não deixou claro quais os requisitos mínimos dos dispositivos onde ADMIN pode ser aplicado
Artigo bem escrito com grande profundidade na parte matemática (Prob. e Estatística)	Não explicou porque escolheu os dados PV, Temperatura e Batimento Cardíaco foram escolhidos
Explicou a escolha dos parâmetros definidos na etapa de testes e resultados	Quantidade média de artigos recentes referenciados: 13 de 29 (após 2014). 10 artigos citados de 2010 ou antes
Faz boas comparações com os trabalhos relacionados	

Obrigado