Assinatura Comportamental e Detecção de Anomalias Utilizando *k-means*

Wagner Senger
Prof. Dr. Lourival Ap. de Góis

Uso de dispositivos

- Áreas da tecnologia geram um crescimento significativo de **recursos computacionais** diferentes.

Uso de dispositivos

- Áreas da tecnologia geram um crescimento significativo de **recursos computacionais** diferentes.

Fonte:http://userscontent2.emaze.com/images/1e3a225d-a645-4b87-bc13-12f2ffd4a0ef/6390d395-bec5-457d-8620-e505c2923cfb.jpg

Uso de dispositivos

- Áreas da tecnologia geram um crescimento significativo de **recursos computacionais** diferentes.

Fonte:http://userscontent2.emaze.com/images/1e3a225d-a645-4b87-bc13-12f2ffd4a0ef/6390d395-bec5-457d-8620-e505c2923cfb.jpg

Fonte:http://www.alcidesmaya.com.br/blog/wp-content/uploads/2016/03/redes-de-computadores.jpg

Uso de dispositivos

- Áreas da tecnologia geram um crescimento significativo de **recursos computacionais** diferentes.

Fonte:http://userscontent2.emaze.com/images/1e3a225d-a645-4b87-bc13-12f2ffd4a0ef/6390d395-bec5-457d-8620-e505c2923cfb.jpg

Fonte:http://www.alcidesmaya.com.br/blog/wp-content/uploads/2016/03/redes-de-computadores.jpg

Fonte:https://www.diegomacedo.com.br/wp-content/uploads/2017/06/cloud-computing.png

Uso de dispositivos

- Áreas da tecnologia geram um crescimento significativo de **recursos computacionais** diferentes.

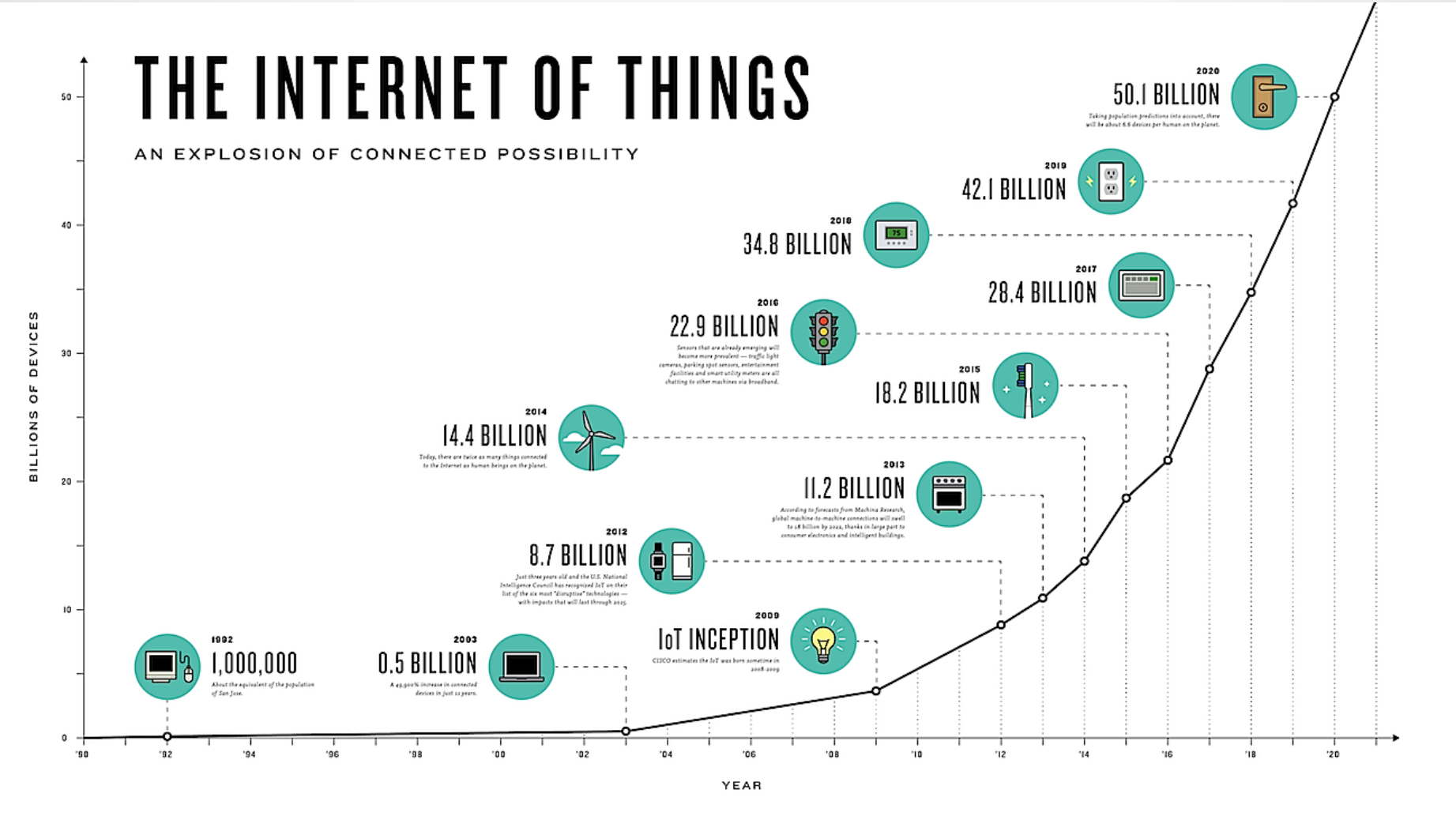
Fonte:http://userscontent2.emaze.com/images/1e3a225d-a645-4b87-bc13-12f2ffd4a0ef/6390d395-bec5-457d-8620-e505c2923cfb.jpg

Fonte:http://www.alcidesmaya.com.br/blog/wp-content/uploads/2016/03/redes-de-computadores.jpg

Fonte:https://www.diegomacedo.com.br/wp-content/uploads/2017/06/cloud-computing.png

Fonte:http://www.crmg.com.br/painel/dbanexos/dbimagens/08-2015/93b6cb553265222bf9dedf7340c8fd65.jpg

Uso de dispositivos



Fonte: http://theconnectivist-img.s3.amazonaws.com/wp-content/uploads/2014/05/Unknown.png

Controle de Recursos

- Recursos conectados a sistemas computacionais podem ser provenientes de ambientes variados.
- Uma falha pode ocasionar problemas de diferentes gravidades, ex:
  - Interromper um vídeo no celular;
  - Lentidão na disponibilização de uma informação;
  - Falhar pode derrubar um serviço web;
  - Deixar de funcionar o freio de um veículo.

É necessário conhecer o recurso

- Não basta monitorar o uso atual do recurso;
- Para identificar se um recurso está dentro do seu funcionamento natural é necessário conhecê-lo.
- O comportamento de um recurso pode ser descrito em forma gráfica, de modo a identificá-lo;
- Assim como uma assinatura pode identificar uma pessoa, é possível criar uma assinatura que identifique um recurso.
- É importante identificar o **padrão** de funcionamento do recurso.

Trabalhos Relacionados

- Estudo: Digital Signature of Network Segment for Healthcare Environments Support. (Proença et al, 2014)
- Algoritmos:
  - ACO (*Ant Colony Optimization*);
  - HW (*Holt Winters*);
  - PCA (*Principal Component Analysis*)

- PCA demonstrou um melhor resultado.

Trabalhos Relacionados

- Estudo: Digital Signature to Help Network Management using flow analysis. (Proença et al, 2015)
- Algoritmos:
  - ACO (*Ant Colony Optimization*);
  - HW (*Holt Winters*);
  - PCA (*Principal Component Analysis*)

- Os algoritmos tiveram resultados aproximados, não sendo elencado um melhor resultado dentre eles.

Trabalhos Relacionados

- Estudo: Anomaly detection system using resource pattern learning (Ohno et al, 2009)
- Algoritmo:
  - *K-means*
- Modelo Auxiliar:
  - HMM (*Hidden Markov Model*)

- Obteve resultados satisfatórios aplicando em conjunto o algoritmo e modelo.

Trabalhos Relacionados

- Estudo: Semantic aware online detection of resource anomalies on the cloud (Bhattacharyya, Jandaghi e Sotiriadis, 2016)
- Metodologia:
  - RNN (*Recurrent Neural Networks*)
- Efetuou geração de padrões de comportamento armazenando os dadosem bancos de dados NoSQL, como MongoDB e Cassandra.

Trabalhos Relacionados

- Estudo: Agent-enabled anomaly detection in resource constrained wireless sensor networks (Usman, 2015)
- Metodologia
  - Criação de um módulo de software acoplado ao SO.
- É apontado no estudo um problema relativo a disponibilidade de recursos.

Assinatura Comportamental

- Propósitos:
  - Permitir conhecer o comportamento de um recurso.
  - Proporcionar a identificação de uma anomalia.
  - Aumentar o nível de controle e segurança em um ambiente.
  - Prever como o recurso se comportará futuramente.
  - Proporcionar um modelo simples e leve para determinar o comportamento do recurso.

Assinatura Comportamental

- Recursos Computacionais possuem um funcionamento padrão dentro de uma semana normal.
- Sequência de funcionamento:

  1 – Monitoramento e armazenamento

  2 – Processamento dos dados e aprendizagem

  3 – Criação da Assinatura do recurso

  4 – Análise do uso de um recurso

Assinatura Comportamental

- Recursos Computacionais possuem um funcionamento padrão dentro de uma semana normal.
- Sequência de funcionamento:

1 – Monitoramento e armazenamento

2 – Processamento dos dados e aprendizagem

3 – Criação da Assinatura do recurso

4 – Análise do uso de um recurso

Assinatura Comportamental

- Pode ser feito internamente ao sistema operacional ou via software ex:
  - Shell script
  - Aplicação desenvolvida com auxílio de LP, ex:
    - Java
    - Python
    - C
- Armazenamento:
  - Em formato texto
  - Diretamente em Banco de dados apropriado.

Assinatura Comportamental

- Recursos Computacionais possuem um funcionamento padrão dentro de uma semana normal.
- Sequência de funcionamento:

1 – Monitoramento e armazenamento

2 – Processamento dos dados e aprendizagem

3 – Criação da Assinatura Comportamental do recurso

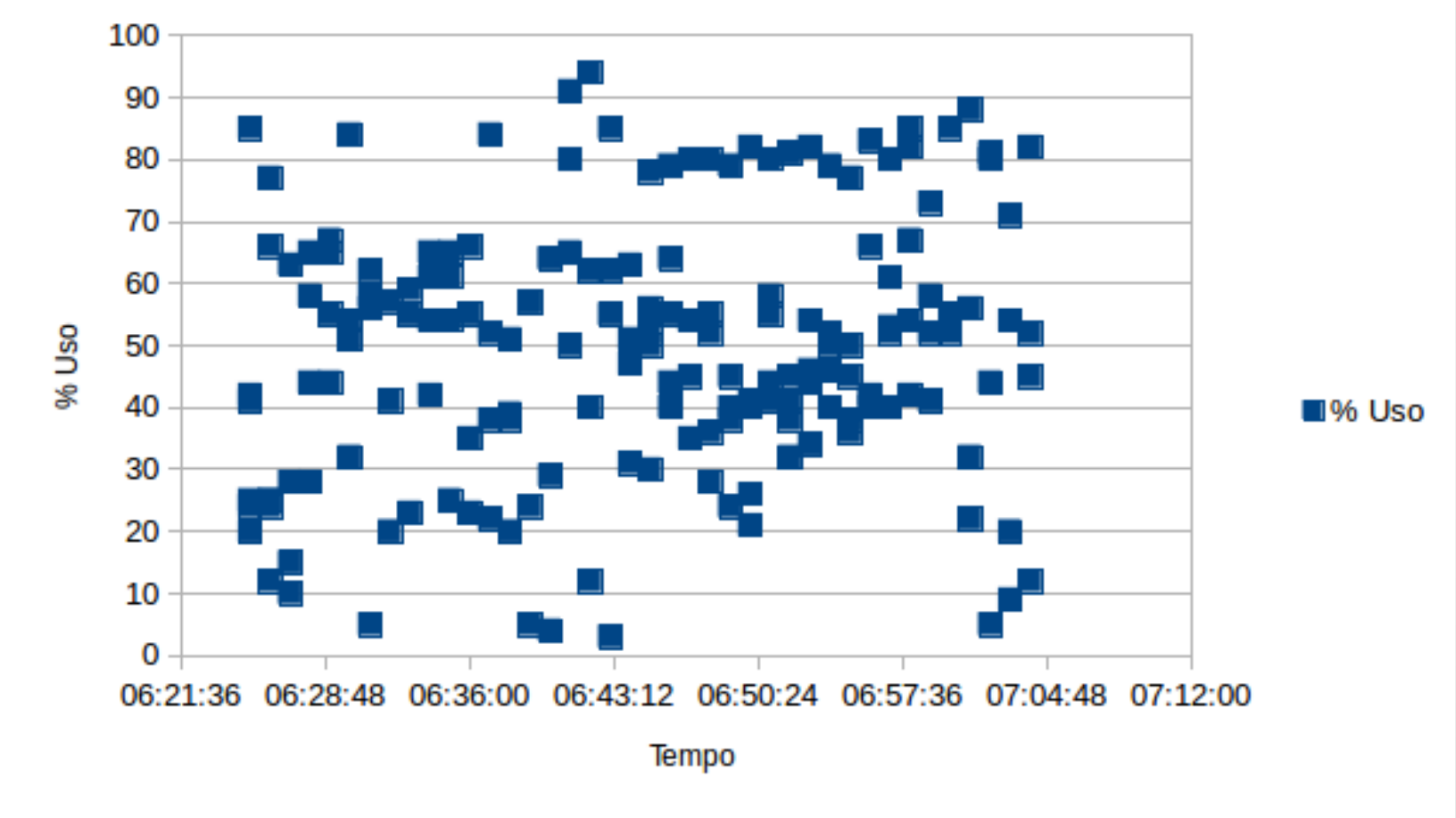
4 – Análise do uso de um recurso

Assinatura Comportamental

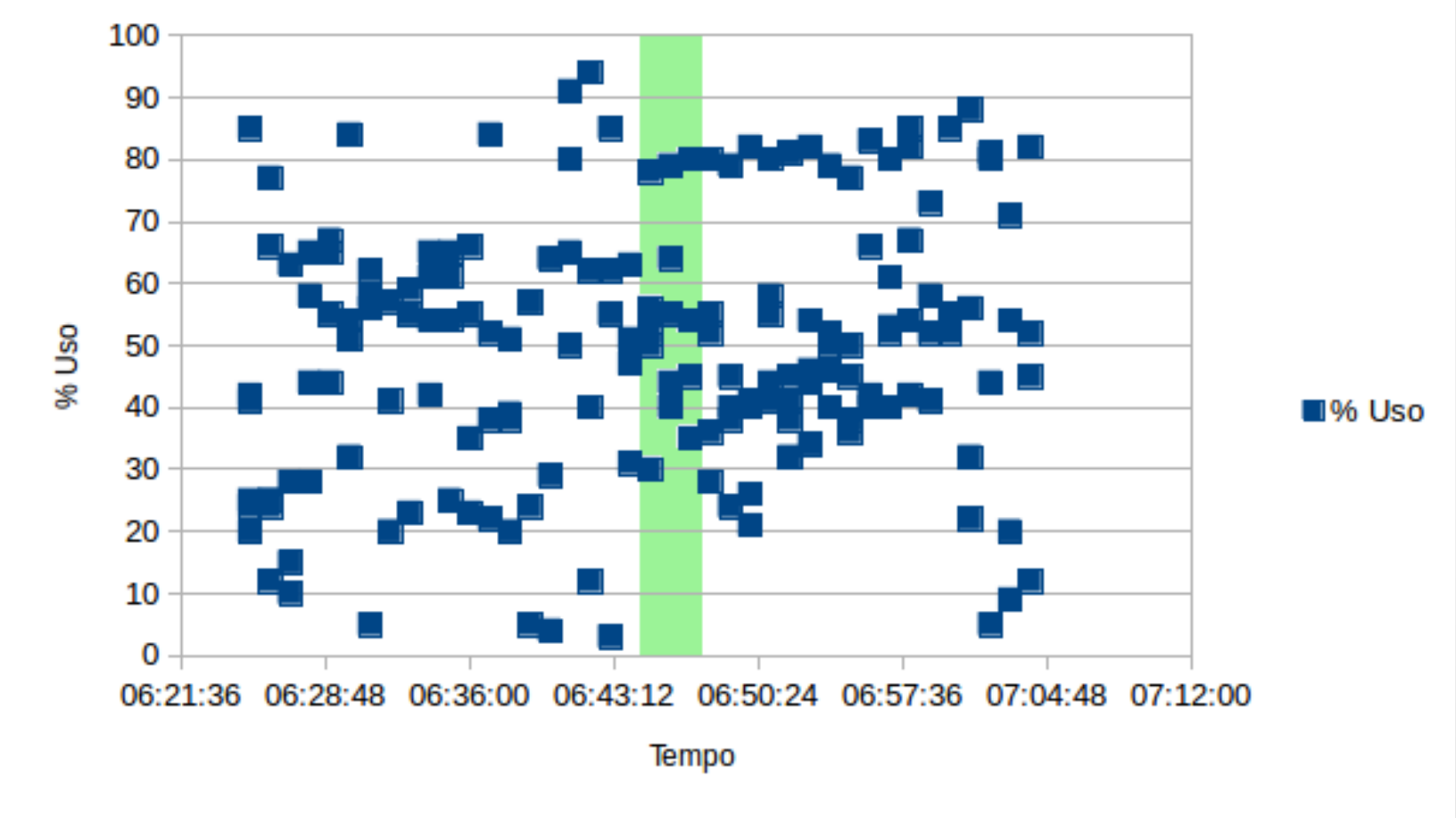
- Utilização do algoritmo de clusterização *k-means*.
  - Algoritmo de clusterização baseado em semelhanças.
  - Determinação do número de clusters solicitados.

- Separação dos dados em janelas de tempo:
  - De **11/09/2017 06:25:40** até **11/09/2017 06:25:46**
- Comparação com janela de tempo de outra semana
  - De **18/09/2017 06:25:40** até **18/09/2017 06:25:46**

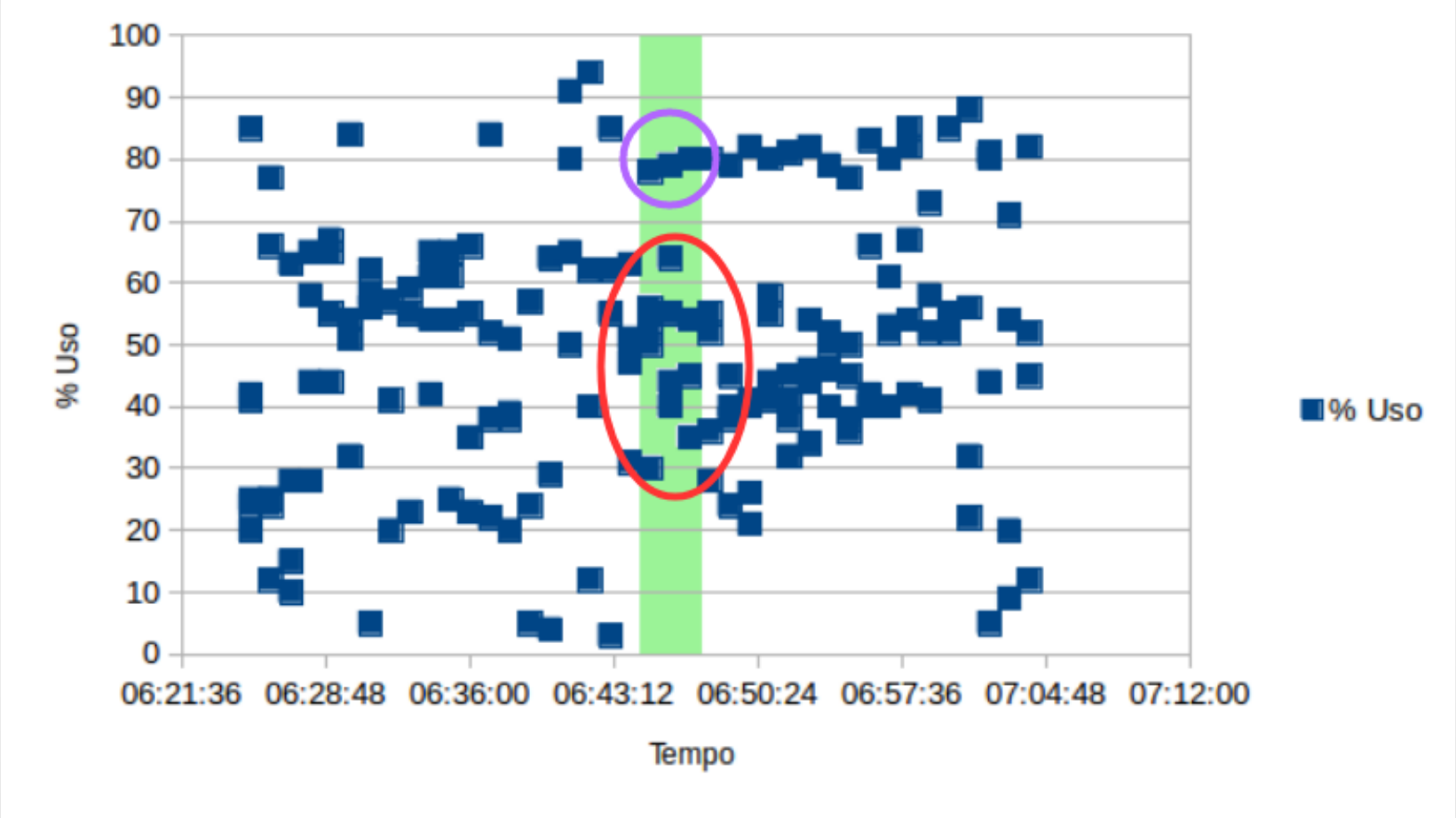
Assinatura Comportamental



Assinatura Comportamental



Assinatura Comportamental



Assinatura Comportamental

- Recursos Computacionais possuem um funcionamento padrão dentro de uma semana normal.
- Sequência de funcionamento:

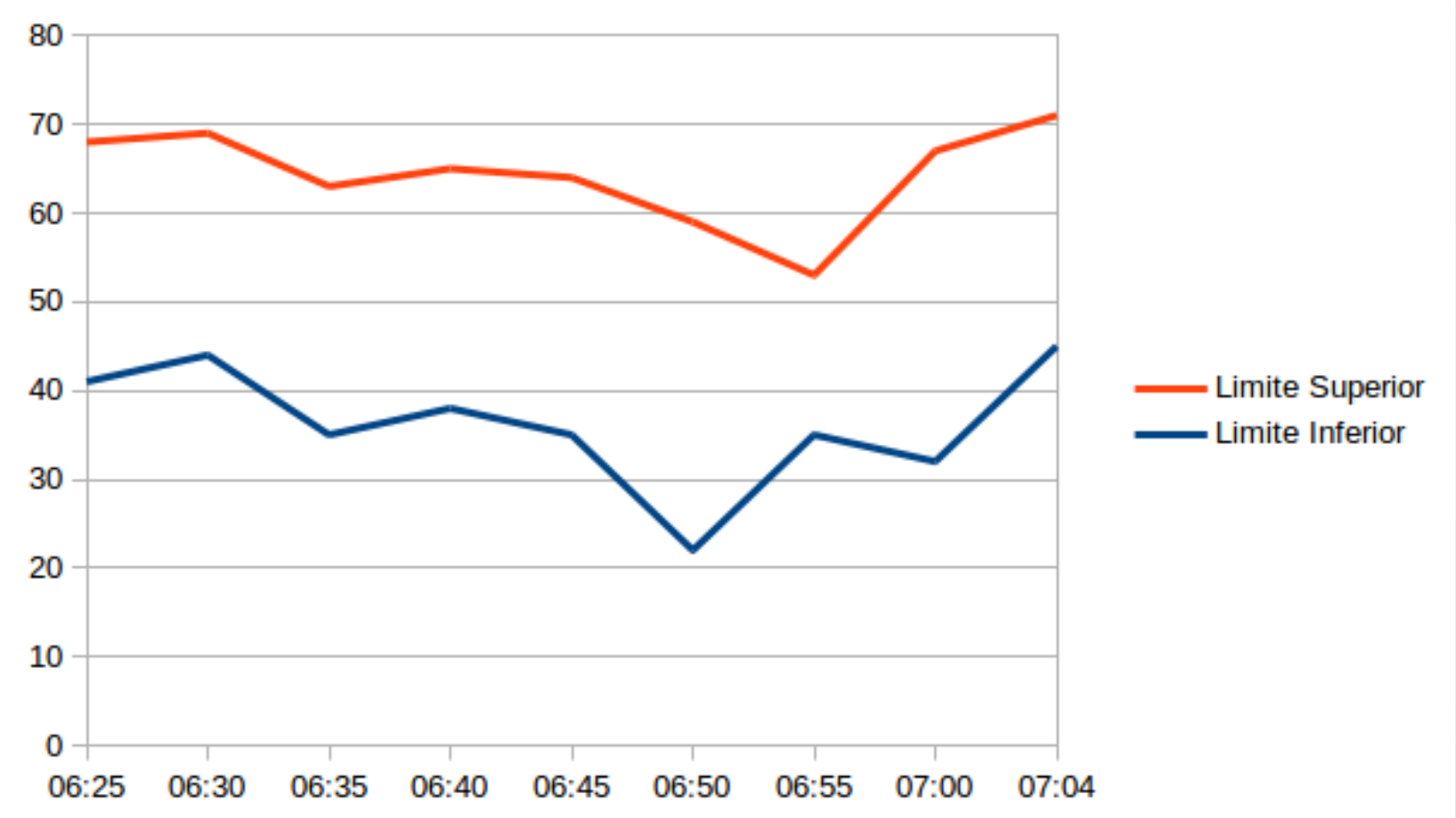
  1 – Monitoramento e armazenamento

  2 – Processamento dos dados e aprendizagem

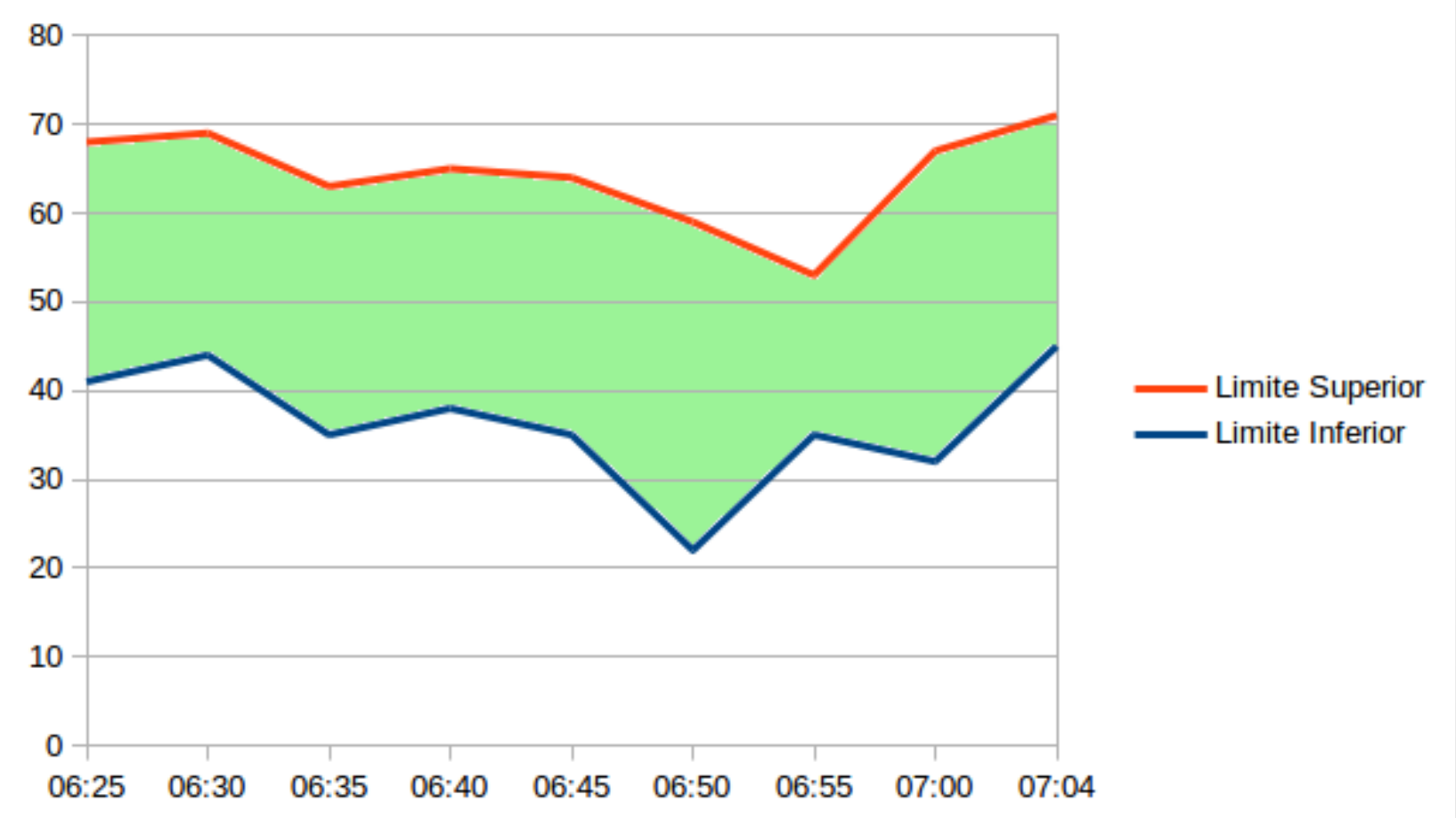
  3 – Criação da Assinatura Comportamental do recurso

  4 – Análise do uso de um recurso

Assinatura Comportamental



Assinatura Comportamental



Assinatura Comportamental

- Recursos Computacionais possuem um funcionamento padrão dentro de uma semana normal.
- Sequência de funcionamento:

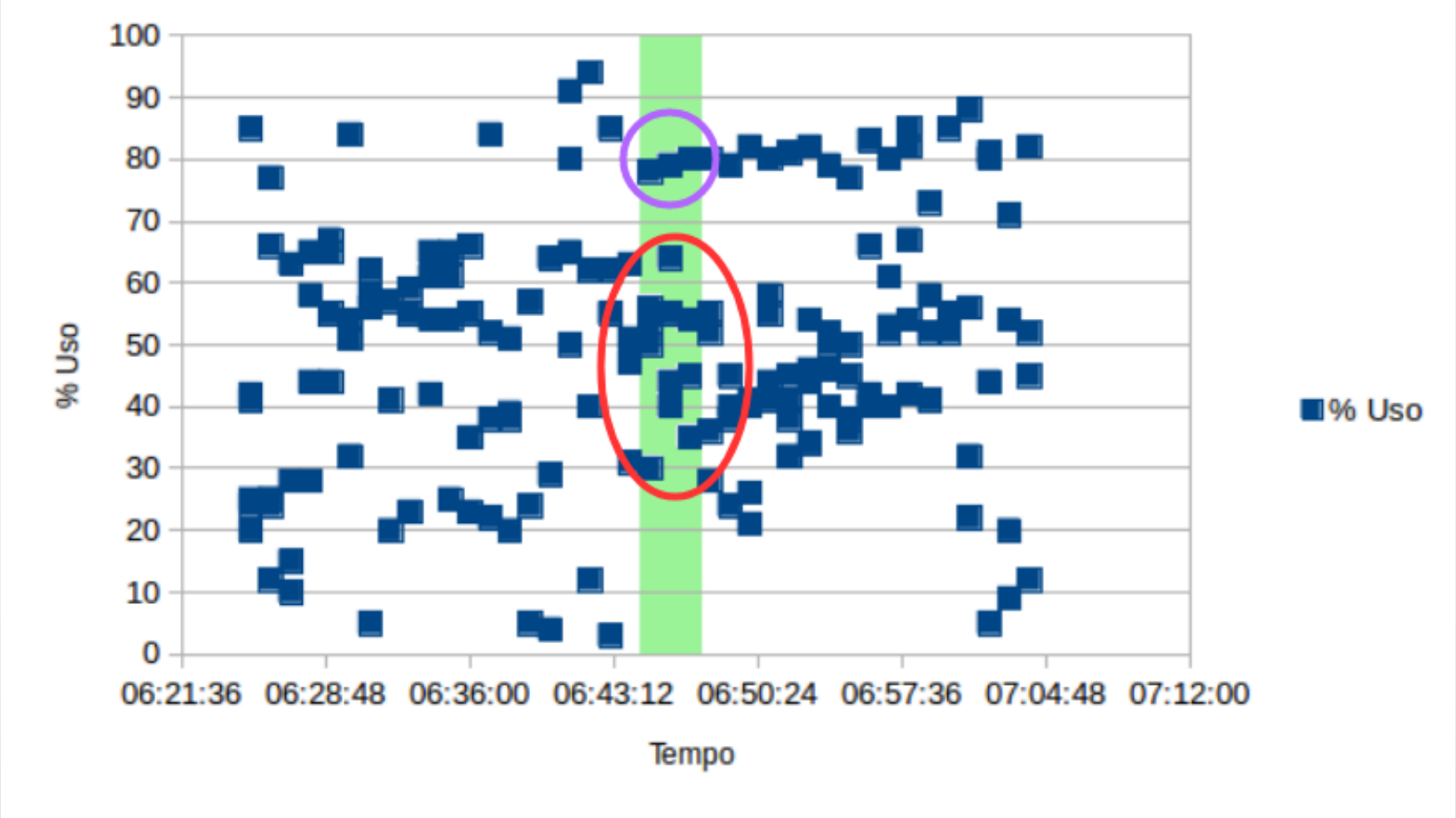
  1 – Monitoramento e armazenamento

  2 – Processamento dos dados e aprendizagem

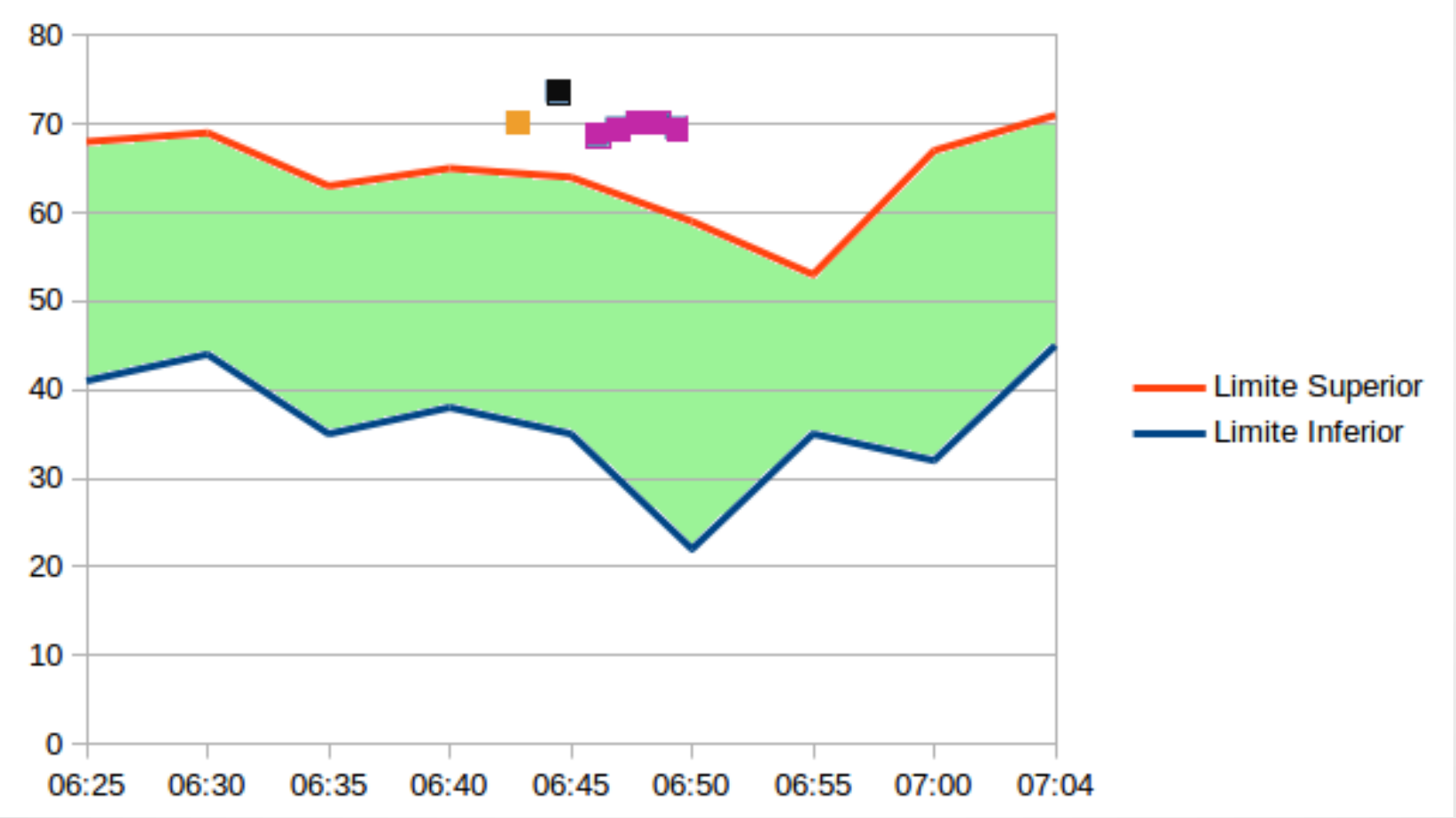
  3 – Criação da Assinatura Comportamental do recurso

  4 – Análise do uso de um recurso

Assinatura Comportamental



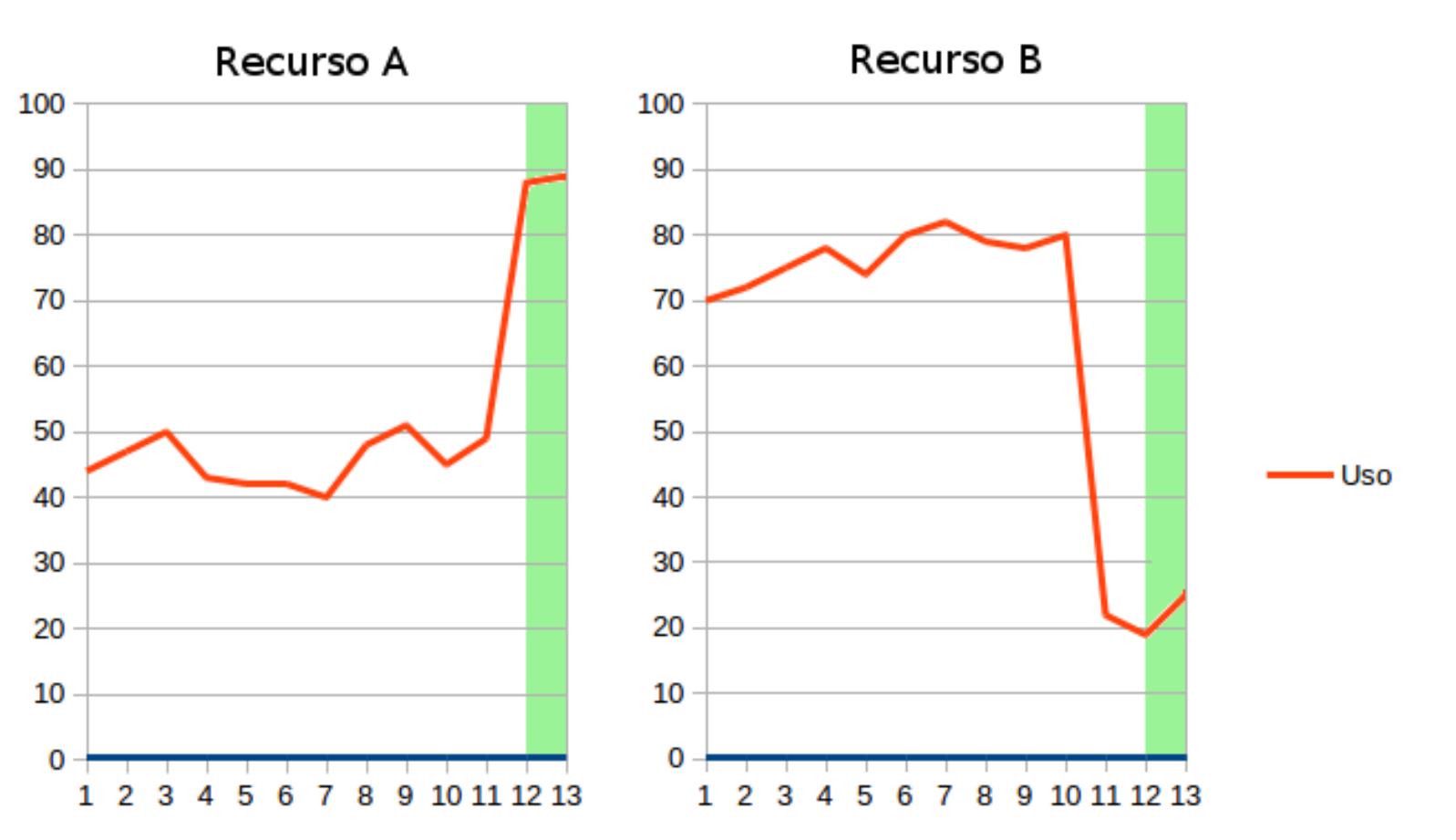
Assinatura Comportamental - *Outliers*



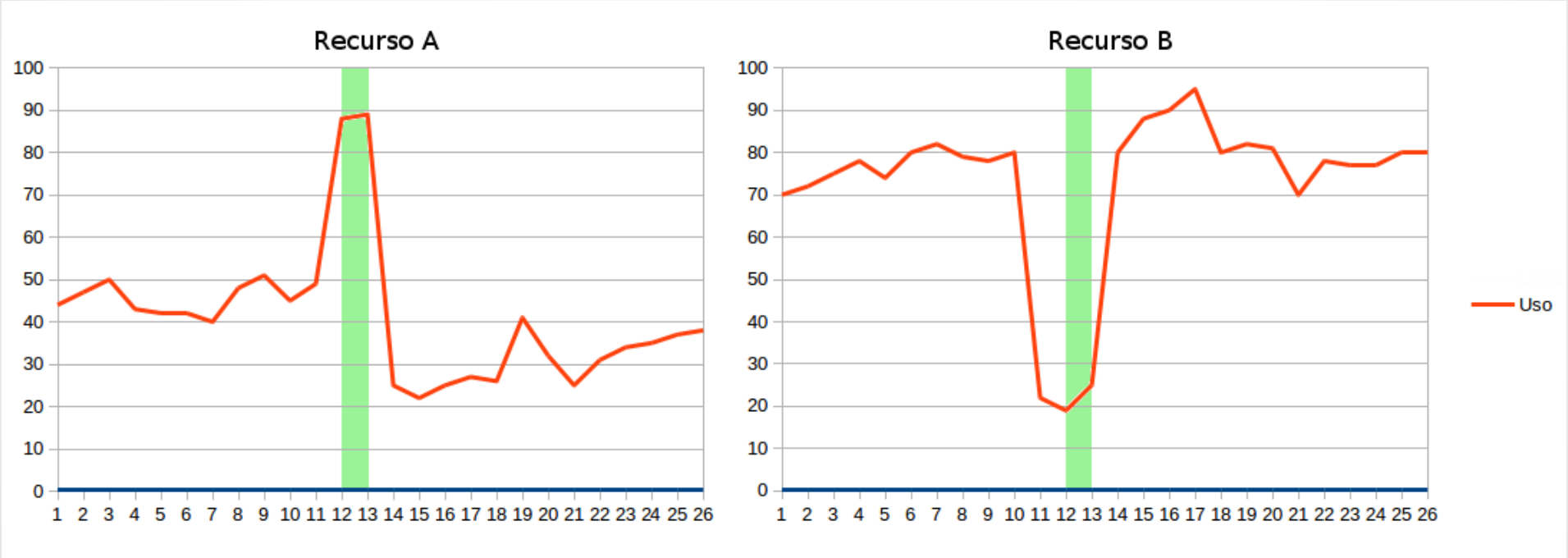
Aplicação em Sistemas Distribuídos

- Distribuição de tarefas para recursos computacionais.
- É necessário conhecer cada recurso para enviar uma solicitação de execução.
- Sem conhecer é possível destinar uma execução para um recurso que irá receber uma grande carga de processamento em um curto tempo;

Aplicação em Sistemas Distribuídos



Aplicação em Sistemas Distribuídos



Obrigado!


PPGCC
Programa de Pós-Graduação
em Ciência da Computação


UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ