



Universidade Federal de Santa Catarina (UFSC)
Programa de Pós-Graduação em Engenharia de Automação e Sistemas (PPGEAS)
Disciplina: Sistemas de Tempo Real II

Análise Temporal Probabilística Baseada em Medições (ATPBM)

Karila Palma Silva
karila.palma@posgrad.ufsc.br

Luís Fernando Arcaro
luis.arcaro@posgrad.ufsc.br

Rômulo Silva de Oliveira
romulo.deoliveira@ufsc.br

Novembro de 2017

1/55

Introdução

3/55

Conteúdo

- 1 Introdução
- 2 Hardware
- 3 Caminhos de execução
- 4 Dados de entrada
- 5 Análise
- 6 Aleatorização temporal
- 7 Exemplo
- 8 Confiabilidade e Aperto
- 9 CRPS

2/55

Introdução

Worst-Case Execution Times (WCETs):

- Sistemas de Tempo Real (STRs) são sistemas computacionais que precisam atender a requisitos de natureza temporal [1].
- Os requisitos temporais aos quais STRs estão sujeitos são expressos em termos dos prazos (*deadlines*) de suas tarefas.
- Para STRs críticos é fundamental o oferecimento de **garantias** quanto ao atendimento a esses prazos mesmo no **pior caso**.
- Testes de escalonabilidade são utilizados para prover tais garantias, e baseiam-se em parâmetros temporais:
 - Impostos pelo contexto de operação, como períodos e *deadlines*.
 - Derivados levando em conta o *software* e *hardware* utilizados, como os **tempos de execução no pior caso** (WCETs) das tarefas.

4/55

Introdução

Estimação de *WCETs*:

- Abordagens **estáticas**:
 - Análise conjunta do código e do executável da tarefa e das características temporais do processador utilizado [2].
 - Impõem limitações sobre o *hardware* e o *software* utilizados.
 - Os valores obtidos são seguros, porém podem ser extremamente pessimistas devido a simplificações do modelo do *hardware*.
- Abordagens **baseadas em medições**:
 - A partir da observação dos tempos de execução efetivamente produzidos pela tarefa enquanto executada na plataforma-alvo [3].
 - São frequentemente utilizadas na indústria de STRs não-críticos.
 - O maior tempo observado pode não ser o *WCET*, e portanto margens de segurança adequadas precisam ser definidas.
- Recentemente, através da Análise Temporal Probabilística Baseada em Medições (ATPBM) e métodos estatísticos.

5/55

Introdução

Análise Temporal Probabilística Baseada em Medições (ATPBM):

- Abordagem proposta recentemente para derivar estimativas probabilísticas de *Worst-Case Execution Time (WCET)*.
- Tais estimativas são também conhecidas na literatura como *Probabilistic Worst-Case Execution Times (pWCETs)*.
- São estimativas de *WCET* que são excedidas com probabilidade máxima que é conhecida e extremamente baixa (e.g. 10^{-15}) [3].
- Baseia-se na aplicação da Teoria dos Valores Extremos (TVE) sobre medições dos tempos de execução da tarefa analisada.
- Considerada promissora para a determinação de *pWCETs* para STRs executando em processadores relativamente complexos.
- Mesmo em sistemas críticos (e.g. aviões), a probabilidade de falhas de natureza não-temporal (e.g. estrutural) **não é nula**.
- Podendo o sistema falhar com certa probabilidade não-nula por causas diversas, o custo/esforço adicional para prover garantias absolutas para os requisitos temporais não se justifica [3].

6/55

Introdução

Análise Temporal Probabilística Baseada em Medições (ATPBM):

- A necessidade de estimativas de *WCET seguras*, associada à natureza dos STRs, impõe diversos requisitos à ATPBM.
- Mede-se tempos de execução da tarefa-alvo na plataforma-alvo.
- As condições de medição devem **se igualar** ou **ser pessimistas** em relação ao pior caso esperado no ambiente de operação.
- Amostras devem ser **representativas** com relação às situações extremas esperadas no ambiente em que o sistema irá operar.
- Os tempos medidos são sujeitos aos requisitos da TVE, como *i.i.d.* e aderência dos máximos a modelos de valores extremos.
- Mesmo que todos os requisitos da TVE sejam satisfeitos, os resultados produzidos podem provar-se não confiáveis.
- Condições de medição e características construtivas do sistema analisado impactam diretamente na confiabilidade dos resultados.

7/55

Introdução

Análise Temporal Probabilística Baseada em Medições (ATPBM):

- Tempos de execução diferem fenomenologicamente daqueles para os quais a TVE foi inicialmente criada (e.g. nível do mar).
- Variabilidade temporal é induzida por fatores **controláveis** e/ou **especulativos**, e pode portanto não ter comportamento aleatório.
- As principais fontes de variabilidade temporal são:
 - O **hardware** do processador utilizado.
 - Os **caminhos de execução** que são efetivamente medidos.
- Tanto os efeitos do **hardware** quanto dos **caminhos de execução** dependem direta ou indiretamente dos **dados de entrada** usados.
- É necessário que todas as decisões tomadas no processo de aplicação da ATPBM suportem a confiabilidade do *pWCET*.

8/55

Hardware

9/55

Hardware

Elementos de latência variável:

- Dependente das instruções e dos dados previamente acessados:
 - **Exemplo:** Memórias *cache* têm latência diferente em *hit / miss*, e dados acessados e seus vizinhos permanecem armazenados.
 - Estado inicial nulo [4] **não é necessariamente** o de pior caso.
 - Determinar o estado inicial de pior caso é um problema difícil [2].
 - Aleatorizar o estado inicial durante as medições [5] não garante que o pior caso será efetivamente observado (ele pode ser muito raro).
- Dependente dos dados sobre os quais operam:
 - **Exemplo:** Operações de multiplicação e divisão em ULAs e UPFs podem ser concluídas mais rapidamente para valores baixos.
 - Pode-se forçar latência máxima durante as medições [6].
- Dependente do comportamento de outros elementos:
 - **Exemplo:** Barramentos compartilhados exigem acesso exclusivo, portanto atrasos podem ser necessários para arbitragem.
 - Isolamento temporal [7] pode eliminar dependência.
 - Medição sob interferência máxima:
 - Necessárias abordagens específicas para cada elemento.
 - Determinação da condição de pior caso pode ser difícil.

11/55

Hardware

Hardware:

- Tarefas de *software* apresentam diferentes tempos de execução quando executadas sobre diferentes plataformas de *hardware*.
- Tempos de execução são afetados pelas **latências** dos elementos internos do processador que realizam a execução de instruções.
- Elementos de entrada e saída de dados também podem influenciar os tempos de execução, direta ou indiretamente.
- A introdução de elementos de aceleração torna os tempos mais baixos, porém **variáveis** e dependentes do **histórico de execução**.
- Portanto, múltiplas execuções de uma tarefa podem produzir tempos **diferentes** entre si devido a características do *hardware*.
- No contexto da ATPBM essa variabilidade inerente induz problemas de **representatividade amostral**, uma vez que:
 - A pior combinação de latências não ocorre frequentemente.
 - Determinar condições para forçar sua ocorrência é difícil.

10/55

Caminhos de execução

12/55

Caminhos de execução

Caminhos de execução:

- São os caminhos percorridos pelo fluxo de controle da tarefa.
- A eficácia do sistema é altamente dependente da forma como os caminhos são selecionados.
- O que define qual **caminho** é executado são:
 - Variáveis de entrada do programa.
 - Variáveis permanentes alteradas em execuções anteriores.
- A quantidade de caminhos que existem pode ser um número tratável ou intratável explicitamente:
 - Tarefa **sem** desvio e **sem** laço: Existe apenas um caminho.
 - Tarefa **com** desvio mas **sem** laço: Depende da combinação dos desvios, mas em geral é tratável explicitamente.
 - Tarefa **sem** desvio mas **com** laço: Depende da combinação dos laços, mas geralmente é tratável explicitamente.
 - Tarefa **com** desvio e **com** laço: Depende da combinação dos laços e dos desvios, sendo em geral intratáveis explicitamente.

Caminhos de execução

Caminhos de execução impossíveis:

- Alguns caminhos podem nunca ser executados, por serem:
 - Impossíveis pela semântica do programa, por exemplo devido a condições contraditórias consecutivas.
 - Impossíveis pela semântica do ambiente, por exemplo devido a entradas impossíveis de acontecer no ambiente real de operação.

Exemplo: Caminho impossível pela semântica do programa.

```

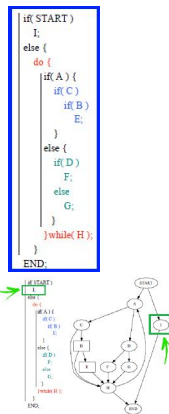
1: void task (float x, y) {
2:   for(int i = 1; i <= 100; i++) {
3:     if (x > 10)
4:       x = x * 2;
5:     else
6:       x = x + 2;
7:     if (x < 0)
8:       y = x;
9:   }
10:}
    
```

O caminho 2-3-4-7-8 é impossível, pois se $x > 10$ então $x * 2 > 0$.

Caminhos de execução

Problema em aberto na ATPBM:

- Medição da **execução completa**: Do início ao fim.
 - Em operação, quais são os caminhos executados?
 - Em operação, com que frequência são executados?
 - Qual(is) caminho(s) de execução é necessário medir?
 - Quantas vezes é necessário medir cada caminho para obter representatividade amostral?
- Medição de **segmentos**: De partes da tarefa.
 - Exemplo: Blocos básicos, i.e. seqüências de instruções sem desvios (um ponto de entrada e um de saída).
 - Como combinar as medições das partes?

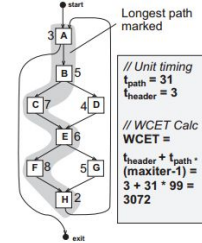


O desafio está em encontrar o(s) pior(es) caminho(s) e garantir que este(s) tenha(m) sido executado(s) representativamente.

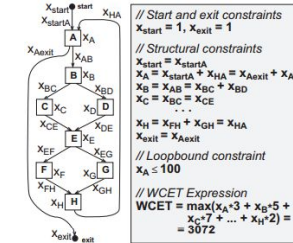
Caminhos de execução

- Existem três classes principais de métodos que combinam tempos medidos ou determinados analiticamente [2]:

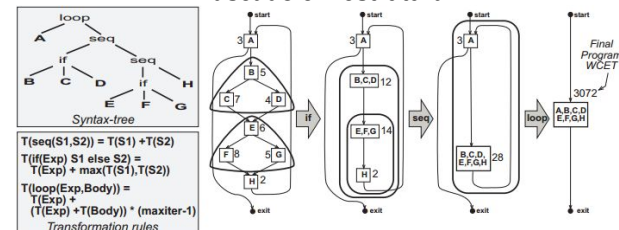
Baseado em caminho



IPET



Baseado em estrutura



Caminhos de execução

Soluções de análise estática [2]:

- **Análise de valor** para a detecção de caminhos impossíveis.
 - Pode concluir que um desvio nunca será tomado, pois o registrador utilizado para avaliação sempre terá um valor equivalente a falso.
 - Tal caminho poderá ser seguramente desconsiderado na busca pelo pior caminho de execução.
- **IPET** para busca de piores caminhos de tempo de execução.
 - Foi apresentada como uma técnica eficiente de busca de caminhos.
 - Apresenta estimativas precisas e independência de dados de entrada do programa analisado.
- **Execução simbólica** para geração de dados de entrada para executar um caminho específico.
 - O programa é executado com valores simbólicos.
 - É produzido um predicado de caminho que determina os intervalos das variáveis de entrada que ocasionam a execução do caminho.

No contexto da ATPBM a determinação do(s) pior(es) caminho(s) de execução é considerada um problema em aberto da área.

17/55

Dados de entrada

18/55

Dados de entrada

Dados de entrada:

- Considera-se dados de entrada as variáveis:
 - Passadas como parâmetro para a tarefa a cada execução.
 - Cujo valor é mantido entre diferentes execuções da tarefa (estado).
- Têm influência nos tempos de execução tanto através do **hardware** quanto dos **caminhos de execução**.

No contexto da ATPBM:

- Os dados de entrada utilizados durante a obtenção das medições é uma decisão crítica para a confiabilidade dos resultados.
- Precisam garantir representatividade amostral com relação aos maiores tempos de execução possíveis.
- Medir todas as combinações possíveis de entradas é **impraticável**.
- Utiliza-se dados de entrada **fixos** ou escolhidos **aleatoriamente**.

19/55

Dados de entrada

Dados de entrada fixos:

- Fixa-se os dados de entrada em valores que acredita-se executar o(s) caminho(s) que produz(em) os **maiores tempos de execução**.
- As entradas de pior caso são em geral desconhecidas, e sua determinação é um problema extremamente difícil [2, 8].
- Como determinar os conjuntos de valores a serem fixados?
- Com que frequência utilizar cada conjunto durante as medições?

Dados de entrada aleatórios:

- Não tem bom desempenho em termos de cobertura.
- Estruturas que são apenas executadas com baixa probabilidade muitas vezes não são cobertas.
- O caminho de pior caso pode ser extremamente raro e as condições para que ele seja executado desconhecidas.
- Quantas medições são necessárias para garantir representatividade?

20/55

Considerações

Análise Temporal Probabilística Baseada em Medições (ATPBM):

- Na ATPBM é necessário que a amostra de tempos de execução seja representativa ou pessimista em relação ao pior cenário para as fontes de variação.
- Caso contrário, os resultados produzidos podem ser invalidados caso interferências maiores que as testemunhadas na medição sejam possíveis no ambiente real de operação do sistema.
- Conforme apresentado, as principais fontes de variação temporal dependem direta ou indiretamente dos dados de entrada e são:
 - O *hardware* do processador utilizado.
 - Os *caminhos de execução* que são efetivamente medidos.
- O pior cenário geralmente é desconhecido e difícil de ser determinado, algumas restrições são assumidas.
- A partir de uma amostra de tempos de execução representativa, é possível utilizar ATPBM e TVE para obter estimativas *pWCETs*.

21/55

Análise

22/55

Análise

Para a aplicabilidade da TVE na ATPBM devemos lembrar que:

- Existem requisitos básicos que devem ser satisfeitos:
 - O fenômeno analisado deve poder ser considerado *aleatório*, de forma que faça sentido a aplicação de teorias probabilísticas.
 - As observações precisam ser modeláveis através de variáveis aleatórias *independentes e identicamente distribuídas* (i.i.d.).
 - Aderência a um dos *modelos probabilísticos* utilizados pela TVE.
- Esses requisitos podem ser verificados através de testes de hipóteses.
- Existem decisões de análise que ainda não foram definidas:
 - Tamanho da amostra utilizado.
 - Tamanho dos blocos.
 - Modelo probabilístico utilizado.

Essas decisões são críticas, pois precisam garantir representatividade. Existem apenas algumas especulações para sua definição.

23/55

Análise

Tamanho da amostra:

- Pequenas amostras dão confiança limitada em relação à população. E grandes amostras têm alto custo de coleta.
- A amostra precisa conter eventos temporais relativamente raros [9].
- Em [10] é proposta uma abordagem na qual $N + N_{delta}$ medições são coletadas e mais N_{delta} são obtidas a cada iteração, até que testes apropriados de ajuste dos modelos convirja.

- A convergência é detectada nessa abordagem através do *Continuous Ranked Probability Score (CRPS)*, calculado como:

$$\sum_{i=0}^{+\infty} (f_x(i) - f_y(i))^2$$

onde f_x e f_y são as funções de distribuição de dois modelos estatísticos consecutivos ajustados a amostras de tamanho crescente a serem comparados.

- Quando o *CRPS* mantêm-se abaixo de um limiar (usa-se 0.01 em [10] e 0.001 em [11]) durante um número definido de iterações (usa-se 5 em [10, 11]), considera-se que o modelo convergiu.

24/55

Análise

Tamanho dos blocos:

- Os efeitos do uso de diferentes tamanhos de blocos foram avaliados em alguns trabalhos [10, 12, 13].
- Porém, ainda não existem conclusões consistentes com relação aos melhores valores a serem utilizados.
- O tamanho de bloco que vem sendo mais utilizado nos trabalhos da literatura é 50 [10].

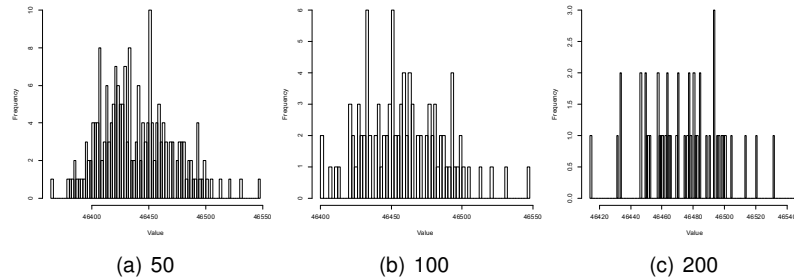


Figura: Tamanho de bloco

Análise

Modelo probabilístico:

- Alguns autores defendem o emprego da distribuição de Gumbel na análise de tempos de execução [14, 10].
- Outros sugerem empregar *Generalized Extreme Value (GEV)* como meio de manipular de forma adequada os valores máximos que convergem para distribuições *Weibull* ou *Fréchet* [15].
- *GEV* é capaz de apresentar o mesmo comportamento que as distribuições *Weibull*, *Gumbel* ou *Fréchet* dependendo do *shape*.
- Os intervalos de confiança do *shape* também permitem expressar incerteza em relação ao qual, entre os três modelos da *GEV*, melhor representa os dados analisados.
- O processo de estimar o valor do *shape* torna-se uma parte crítica do processo de ajuste do modelo, pois pequenas variações de seu valor podem levar a estimativas *pWCET* muito diferentes [16].

Análise

Modelo probabilístico:

↳ *GEV*:

- Seus parâmetros: *location* (μ), *scale* (σ) e *shape* (ξ) têm efeito diretamente proporcional no *pWCET*.
- Efeito do *shape* é mais expressivo do que *location* e *scale*.
- Se $\xi \leq 0$ a cauda direita é “curta” (decréscimo no mínimo exponencial), e estimativas serão relativamente baixas em comparação com os tempos médios.
- Se $\xi > 0$ a cauda direita é longa, e as estimativas poderão tornar-se altas a ponto de serem **inúteis para finalidades práticas**.
- Exemplo: Para uma tarefa que demora em média 10^5 ciclos de *clock* para executar, encontra-se um *pWCET* de 10^{10} ciclos de *clock*. A estimativa pode ser segura, mas a reserva de recursos é inaceitável na prática.

Análise

Modelo probabilístico:

↳ *Gumbel*:

- Cauda direita exponencial ($\xi = 0$).
- Conforme apresentado em [13], o modelo de *Gumbel* parece ser mais consistente e estável em relação às estimativas fornecidas.
- Se $\xi < 0$ (*Weibull*), o modelo de *Gumbel* apresenta algum pessimismo, mas pode ser usado com confiança razoável na confiabilidade dos resultados obtidos.
- Se $\xi > 0$ (*Fréchet*), o modelo de *Gumbel* produz estimativas não confiáveis, uma vez que não é aplicável a tais cenários.

Considerações

Análise Temporal Probabilística Baseada em Medições (ATPBM):

- A aplicabilidade da TVE na ATPBM para obter estimativas $pWCETs$ nem sempre é possível.
- Plataformas de *hardware* que tendem a gerar tempos discretos dificultam ou impossibilitam a análise.
- Se os requisitos básicos de aplicabilidade não são satisfeitos e a amostra de tempos de execução utilizada não é representativa, não se pode ter credibilidade dos resultados da análise.
- A analisabilidade através da TVE tem sido favorecida com a aleatorização temporal de *hardware*, porém não é garantida.

29 / 55

Aleatorização temporal

30 / 55

Aleatorização temporal

Aleatorização temporal:

- *Hardware* tradicional:
 - É regido por regras determinísticas, e não probabilísticas.
 - Tende a gerar tempos de execução que variam de forma incompatível com modelos probabilísticos típicos.
- *Hardware* temporalmente aleatorizado [6]:
 - Decisões tipicamente especulativas ou dependentes do histórico de execução passam a ser tomadas de forma (pseudo-)aleatória.
 - Tempos de execução patológicos tendem a tornar-se raros, pois passam a depender de cadeias longas de eventos “aleatórios” [3].
 - Introduce comportamento probabilístico, melhorando assim a analisabilidade dos tempos através de ferramentas estatísticas.
 - Torna a natureza dos sistemas computacionais mais próxima dos fenômenos para os quais a TVE foi inicialmente desenvolvida.
 - **Não há garantias** quanto à analisabilidade dos tempos através da TVE, portanto testes de aplicabilidade precisam ser usados [15].

31 / 55

Exemplo

32 / 55

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Tarefa:

- A tarefa escolhida para realização da análise foi a **Bsort**, que ordena um vetor de 10 elementos pelo método *bubble sort*.

↳ Caminho de execução:

- **Fixamos os dados de entrada** que acredita-se executar o caminho de execução mais longo.
- O caminho que executa o maior número possível de operações elementares, a fim de induzir altos tempos de execução.
- Usamos vetores de números inteiros em ordem inversa.

33/55

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Hardware:

- Os tempos foram medidos em uma plataforma de *hardware temporamente aleatorizada*, executada a 50MHz em *FPGA*.
- **Processador dual-core:**
 - Memórias *RAM* separadas (instruções e dados).
 - Acessadas através de um barramento compartilhado arbitrado com política aleatorizada.
- **Pipeline** de cinco estágios:
 - Implementa instruções *MIPS*.
 - Produz latências que variam de acordo com seu comportamento interno e do de outros elementos (e.g. memórias *cache*).
- **Memória cache** associativa por conjuntos aleatorizada [17].
- **Gerador aleatório** que randomiza o tempo no nível de *hardware*:
 - Latência de um ciclo de *clock*.
 - Combina resultados de dois geradores independentes e diferentes.

34/55

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Hardware:

- A plataforma utilizada – **DPArptdm**:
 - Não usa memória *cache*.
 - Usa política de arbitragem de barramentos aleatorizada.



35/55

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Medição:

- A tarefa foi executada de forma exclusiva e ininterrupta.
- De acordo com as sugestões de [4, 10, 6]:
 - Configurou-se as unidades lógicas para produzir latência máxima.
 - Operações de multiplicação e divisão inteiras levam sempre 32 ciclos de *clock* para completar.
 - Reiniciamos o estado do *hardware* antes das execuções.
- A variabilidade de tempo decorre do processador randomizado.
- Medidas obtidas em condições quase ideais para ATPBM.
- Não é garantido que os *WCETs* reais sejam observáveis.

36/55

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Tamanho da amostra:

- Utilizou-se uma amostra de tamanho 50000.

↳ Tamanho de bloco:

- Utilizou-se blocos de tamanho 50, como proposto em [10].

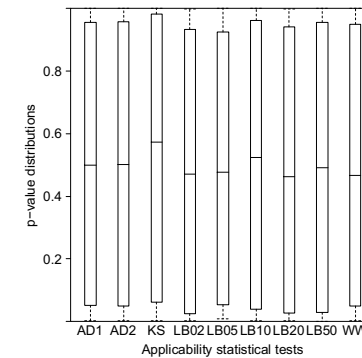
37/55

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Premissas básicas de aplicabilidade:

Testes de i.i.d:



Os resultados dos testes estatísticos são aceitáveis, uma vez que os valores de p são distribuídos quase uniformemente na faixa $[0, 1)$.

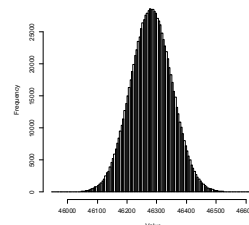
38/55

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

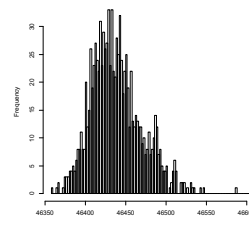
↳ Histograma de dados crus:

- Coletou-se uma amostra composta de 1000 blocos de tamanho 50 de tempos de execução da tarefa analisada:



↳ Histograma de máximos:

- Considerou-se o valor máximo de cada bloco e descartou-se os demais valores, gerando então um conjunto reduzido de máximos de bloco:



39/55

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Ajuste de modelo:

- Ajustou-se os máximos de bloco às distribuições *GEV* e *Gumbel*.
- Para o ajuste dos parâmetros da distribuição utilizou os métodos:
 - *L-moments* – *GEV*.
 - *Maximum Likelihood Estimation (MLE)* – *Gumbel*.

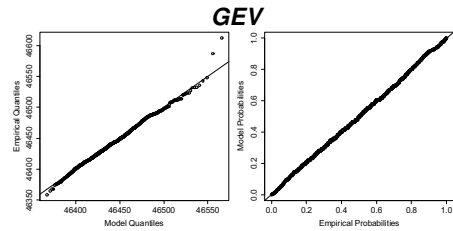
↳ Verificação de ajuste:

- A fim de evidenciar a aderência dos dados às distribuições e, assim, a validade dos resultados, utilizou-se:
 - Gráficos de comparação de quantis.
 - Gráficos de comparação de probabilidades

40/55

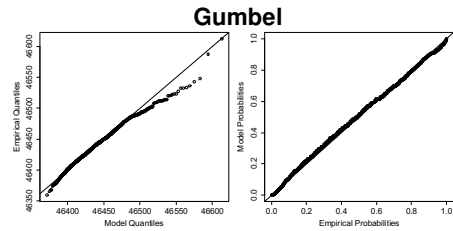
Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):



(a) GEV quantile plot (b) GEV probability plot

- O modelo *GEV* representa corretamente os tempos de execução máximos observados, ressaltando que é suscetível a aderir a uma das distribuições *Weibull*, *Gumbel* ou *Fréchet*.



(c) Gumbel quantile plot (d) Gumbel probability plot

- O modelo *Gumbel* não representa bem os máximos analisados, mas leva a limites superiores confiáveis por superestimar quantis altos.

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Inferência:

- Utiliza-se o modelo probabilístico ajustado aos dados amostrais para realizar inferências.
- Com base em **evidências presentes na amostra** infere-se características **da população**.

↳ Determinação de valores-limite:

- Trataremos especificamente da determinação de **valores-limite** que espera-se que sejam excedidos **com baixa probabilidade**.
- Ou seja, limites para os quais espera-se que a observação de valores maiores no fenômeno analisado seja **suficientemente rara**.
- Utilizaremos para estimar limites probabilísticos para o tempo de execução máximo da tarefa, ou seja, estimar seu *pWCET*.

Exemplo

Análise Temporal Probabilística Baseada em Medições (ATPBM):

↳ Estimativa do pWCET:

Modelo	<i>GEV</i>	<i>Gumbel</i>
μ	46425.6958	46424.2924
σ	27.4849	27.3196
ξ	-0.0934	0
LI IC μ	46423.9115	46422.5036
LS IC μ	46427.5694	46426.0811
LI IC σ	26.0689	26.0484
LS IC σ	28.8507	28.5908
LI IC ξ	-0.1390	0
LS IC ξ	-0.0526	0
pWCET(10^{-6})	46638.9927	46801.7270
pWCET(10^{-7})	46654.6620	46864.6327
pWCET(10^{-8})	46667.2991	46927.5385
pWCET(10^{-15})	46708.2801	47367.9007

Confiabilidade e Aperto

Confiabilidade e Aperto

No artigo [13]:

“On Using GEV or Gumbel Models when Applying EVT for Probabilistic WCET Estimation”

Avaliamos os modelos *GEV* e *Gumbel* com relação à confiabilidade e o aperto das estimativas *pWCET* produzidas no contexto da ATPBM.

Confiabilidade

- Comparamos estimativas *pWCET* com probabilidade de excedência 10^{-15} (e seus intervalos de confiança de 95%) com o maior valor observado para uma amostra de tamanho 10^8 .

Aperto

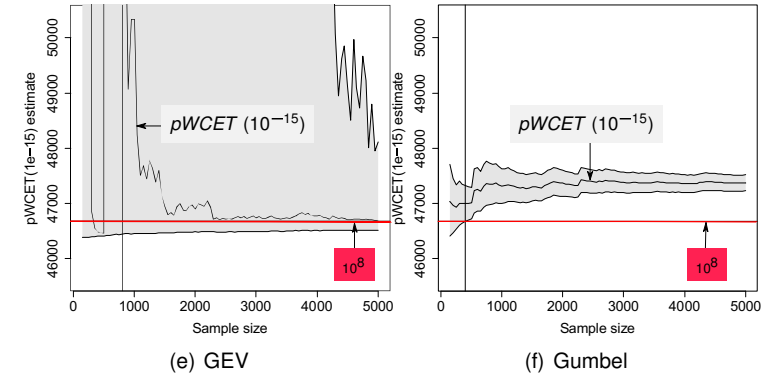
- Analisamos a distância entre os valores máximos observados em amostras de tamanho 10^6 e 10^8 e a estimativa *pWCET* com probabilidade de excedência de magnitude intermediária (10^{-7}).

CRPS

- Como contribuição adicional, avaliamos a métrica CRPS (com limiar 0.001 e 5 iterações) para determinar tamanhos de amostra.

Confiabilidade e Aperto

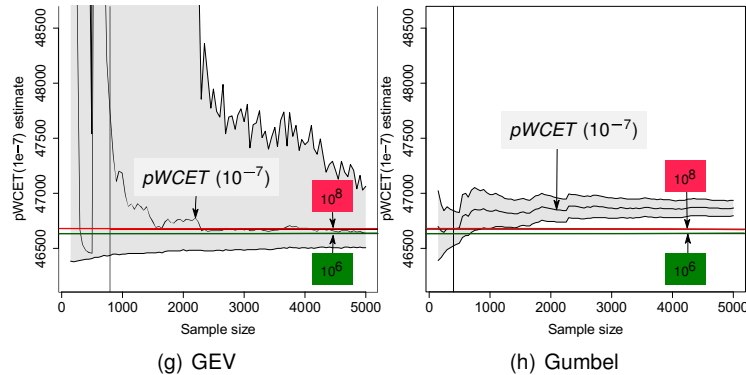
Confiabilidade:



- **GEV:** Deixa dúvidas quanto à confiabilidade – converge para valores muito próximos do maior valor observado em 10^8 .
- **Gumbel:** É potencialmente confiável – todos os *pWCET* (10^{-15}) e seus intervalos de confiança estão acima do valor máximo (10^8).

Confiabilidade e Aperto

Aperto:

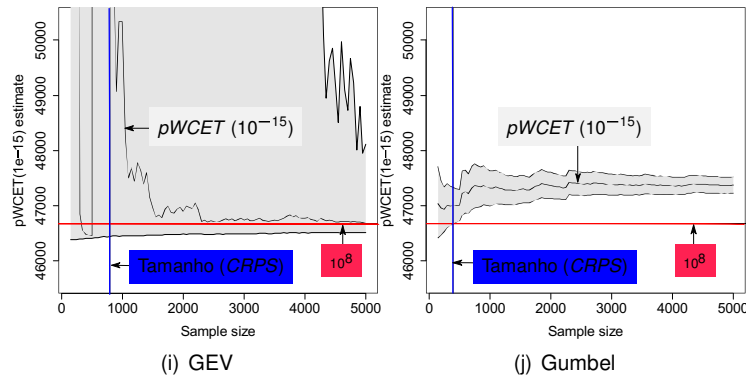


- **GEV:** Converte para valores localizados entre os dois máximos 10^6 e 10^8 , o que sugere que as estimativas são apertadas.
- **Gumbel:** As estimativas e seus intervalos de confiança estão acima dos maiores valores observados (10^6 e 10^8) – o aperto é menor em relação à *GEV*.

CRPS

CRPS

Continuous Ranked Probability Score (CRPS):







- O CRPS não garante que as estimativas se tornem estáveis.
- As estimativas $pWCETs$ produzidas usando CRPS seriam:
 - **GEV**: Muito maiores do que as obtidas após convergência.
 - **Gumbel**: Ligeiramente menor do que as obtidas após convergência




- 📖 J. W.-S. Liu, [Real-Time systems](#), 1st ed. Prentice Hall, 2000.
- 📖 R. Wilhelm, T. Mitra, F. Mueller, I. Puaut, P. Puschner, J. Staschulat, P. Stenström, J. Engblom, A. Ermedahl, N. Holsti, S. Thesing, D. Whalley, G. Bernat, C. Ferdinand, and R. Heckmann, “The Worst-Case Execution-Time Problem - Overview of Methods and Survey of Tools,” [ACM Transactions on Embedded Computing Systems \(TECS\)](#), vol. 7, pp. 36:1–36:53, 2008.
- 📖 F. J. Cazorla, J. Abella, J. Andersson, T. Vardanega, F. Vatrinet, I. Bate, I. Broster, M. Azkarate-askasua, F. Wartel, L. Cucu-Grosjean, F. Cros, G. Farrall, A. Gogonel, A. Gianarro, B. Triquet, C. Hernandez, C. Lo, C. Maxim, D. Morales, E. Quiñones, E. Mezzetti, L. Kosmidis, I. Agirre, M. Fernandez, M. Slijepcevic, P. R. Conmy, and W. Talaboulma, “PROXIMA: Improving Measurement-Based Timing Analysis through Randomisation and Probabilistic Analysis,” in [Euromicro Conference on Digital System Design 2016 \(DSD'16\)](#). IEEE, 2016, pp. 276–285.



- 📖 D. Griffin and A. Burns, “Realism in Statistical Analysis of Worst Case Execution Times,” in [International Workshop on Worst-Case Execution Time Analysis 2010 \(WCET'10\)](#), vol. 15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2010, pp. 44–53.
- 📖 S. M. Petters, “Worst Case Execution Time Estimation for Advanced Processor Architectures,” Ph.D. dissertation, Technische Universität München, 2002.
- 📖 L. Kosmidis, E. Quiñones, J. Abella, T. Vardanega, C. Hernandez, A. Gianarro, I. Broster, and F. J. Cazorla, “Fitting processor architectures for measurement-based probabilistic timing analysis,” [Microprocessors and Microsystems \(MICPRO\)](#), vol. 47B, pp. 287–302, 2016.
- 📖 O. Kotaba, J. Nowotsch, M. Paulitsch, S. M. Petters, and H. Theiling, “Multicore In Real-Time Systems - Temporal Isolation Challenges Due To Shared Resources,” in [Workshop on Industry-Driven Approaches for Cost-Effective Certification of Safety-Critical Mixed-Criticality Systems 2013 \(WICERT 2013\)](#), 2013, p. 6.

-  S. Law and I. Bate, “Achieving Appropriate Test Coverage for Reliable Measurement-Based Timing Analysis,” in [Euromicro Conference on Real-Time Systems 2016 \(ECRTS’16\)](#). IEEE, 2016, pp. 189–199.
-  J. Abella, E. Quiñones, F. Wartel, T. Vardanega, and F. J. Cazorla, “Heart of Gold: Making the Improbable Happen to Increase Confidence in MBPTA,” in [Euromicro Conference on Real-Time Systems 2014 \(ECRTS’14\)](#). IEEE, 2014, pp. 255–265.
-  L. Cucu-Grosjean, L. Santinelli, M. Houston, C. Lo, T. Vardanega, L. Kosmidis, J. Abella, E. Mezzetti, E. Quiñones, and F. J. Cazorla, “Measurement-Based Probabilistic Timing Analysis for Multi-path Programs,” in [Euromicro Conference on Real-Time Systems 2012 \(ECRTS’12\)](#). IEEE, 2012, pp. 91–101.
-  F. Wartel, L. Kosmidis, C. Lo, B. Triquet, E. Quiñones, J. Abella, A. Gogonel, A. Baldovin, E. Mezzetti, L. Cucu-Grosjean, T. Vardanega, and F. J. Cazorla, “Measurement-Based Probabilistic Timing Analysis: Lessons from an Integrated-Modular Avionics Case Study,” in [International Symposium on Industrial Embedded Systems 2013 \(SIES’13\)](#). IEEE, 2013, pp. 241–248.

53/55

-  L. Santinelli, J. Morio, G. Dufour, and D. Jacquemart, “On the Sustainability of the Extreme Value Theory for WCET Estimation,” in [International Workshop on Worst-Case Execution Time Analysis 2014 \(WCET’14\)](#), vol. 39. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014, pp. 21–30.
-  K. P. Silva, L. F. Arcaro, and R. S. de Oliveira, “On Using GEV or Gumbel Models when Applying EVT for Probabilistic WCET Estimation,” in [Real-Time Systems Symposium 2017 \(RTSS’17\)](#). IEEE, 2017, p. 11.
-  J. Hansen, S. Hissam, and G. A. Moreno, “Statistical-Based WCET Estimation and Validation,” in [International Workshop on Worst-Case Execution Time Analysis 2009 \(WCET’09\)](#), vol. 10. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2009, pp. 1–11.
-  G. Lima, D. Dias, and E. Barros, “Extreme Value Theory for Estimating Task Execution Time Bounds: A Careful Look,” in [Euromicro Conference on Real-Time Systems 2016 \(ECRTS’16\)](#). IEEE, 2016, pp. 200–211.

54/55

-  S. G. Coles, [An Introduction to Statistical Modeling of Extreme Values](#), 1st ed., ser. Springer Series in Statistics. Springer, 2001.
-  L. Kosmidis, J. Abella, E. Quiñones, and F. J. Cazorla, “A Cache Design for Probabilistically Analysable Real-time Systems,” in [Design, Automation and Test in Europe Conference and Exhibition 2013 \(DATE’13\)](#). IEEE, 2013, pp. 513–518.

55/55