Empirical Characterization of Discretization Error in Gradient-based Algorithms

Jacob Beal
& Jonathan Bachrach, Joshua Horowitz, Dany Qumsiyeh

SASO 2008
The Challenge of Composition
Outline

- Gradients
- Discretization Error
- Empirical Model
- Predictive Composition
Gradient

Distance from each device to nearest source

Distance in graph is proxy for real distance
Geometric Program: Bisector
Geometric Program: Bisector
Geometric Program: Bisector
Geometric Program: Channel

Source

Destination

(cf. Butera)
Geometric Program: Channel

Source

Destination

(cf. Butera)
Geometric Program: Channel

(c.f. Butera)
Geometric Program: Channel

(cf. Butera)
Geometric Program: Channel

(cf. Butera)
Geometric Program: Channel

Source

Destination

(cf. Butera)
Geometric Program: Channel

Source

Destination

(cf. Butera)
Discretization Error

Prediction: $\varepsilon = \alpha \rho^{-2} d$
Experimental Strategy

- Distribute $n$ devices randomly in area $A$, communicating in $r$ range, for density $\rho$
- Perfect range information, no failures
- Survey wide range of parameters
  - 100 trials/combination, $\sim20K$ total
Four Domains of Behavior
Four Domains of Behavior

Planar source, n=126, r=10

Planar source, n=316, r=10

Planar source, n=631, r=10

Planar source, n=2510, r=10
Density affects error monotonically
Making an Empirical Model

\[ \varepsilon_G = \alpha d + \beta d^{-\gamma} \]

\[ \varepsilon_G = \alpha_1 \rho^{\alpha_2} d + \beta_1 \rho^{\beta_2} d^{(\gamma_1 + \gamma_2 \rho^{\gamma_3})} \]

\[ \sigma_{\varepsilon_G} = \kappa + \lambda d^{-\mu} \]

\[ \sigma_{\varepsilon_G} = \kappa_1 \rho^{\kappa_2} + \lambda_1 \rho^{\lambda_2} d^{(\mu_1 + \mu_2 \rho^{\mu_3})} \]

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>95% confidence bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_1 )</td>
<td>7.8</td>
<td>(6.8, 8.7)</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>-2.14</td>
<td>(-2.19, -2.10)</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>11.2</td>
<td>(10.8, 11.5)</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>-0.516</td>
<td>(-0.526, -0.505)</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>-0.292</td>
<td>(-0.303, -0.282)</td>
</tr>
<tr>
<td>( \gamma_2 )</td>
<td>1.6</td>
<td>(1.3, 1.9)</td>
</tr>
<tr>
<td>( \gamma_3 )</td>
<td>-0.77</td>
<td>(-0.86, -0.69)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>95% confidence bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa_1 )</td>
<td>-25000</td>
<td>(-52000, 2000)</td>
</tr>
<tr>
<td>( \kappa_2 )</td>
<td>-4.5</td>
<td>(-4.9, -4.0)</td>
</tr>
<tr>
<td>( \lambda_1 )</td>
<td>7.40</td>
<td>(7.07, 7.73)</td>
</tr>
<tr>
<td>( \lambda_2 )</td>
<td>-0.529</td>
<td>(-0.541, -0.517)</td>
</tr>
<tr>
<td>( \mu_1 )</td>
<td>-0.278</td>
<td>(-0.283, -0.272)</td>
</tr>
<tr>
<td>( \mu_2 )</td>
<td>11</td>
<td>(5, 16)</td>
</tr>
<tr>
<td>( \mu_3 )</td>
<td>-1.38</td>
<td>(-1.54, -1.21)</td>
</tr>
</tbody>
</table>

Mean

Standard Deviation
Model Fit

Measured vs. Model (Planar Source, \( r=10, \rho=10.5 \))

- Measured
- Mean
- \( \pm 2 \) Std. Dev.

Measured vs. Model (Planar Source, \( r=10, \rho=102.6 \))

- Measured
- Mean
- \( \pm 2 \) Std. Dev.
Source shape matters

density=10.4

density=83.4
Understanding the Transient

Error Structure Tangent to Planar Source ($\theta=209.4$, $r=10$)

- $0 < d \leq 25$
- $25 < d \leq 50$
- $50 < d \leq 75$
- $75 < d \leq 100$
- $100 < d \leq 125$

Tangent Coordinate

Error

Source
Transient Elimination

Point vs. Planar Source, \( n=10k, r=10, \rho=209.4 \)

Coordinate vs. Zero Source, \( n=10k, r=10, \rho=209.4 \)

Point or “true depth” sources eliminate transient
Model Predicts Channel/Bisector

Predicted vs. Measured Bisector Error

Predicted vs. Measured Channel Error
Contributions

- Identified new gradient phenomena
- Created empirical model of gradient error
- Used model to predict channel & bisector error
- Laid foundation for better theoretical prediction of composed gradient programs