Analytical Modeling of Software Development Teams in Globally Distributed Projects

Methods and Tools for Project/Architecture/Risk Management in Globally Distributed Software Development Projects (PARIS’10)

Ricardo M. Czekster, Paulo Fernandes, Afonso Sales, and Thais Webber

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Paleoprospec Project - PUCRS/Petrobras
Funded also by CAPES and CNPq - Brazil
Globally distributed projects

- dispersed multiple sites (borders and time zones)
- difficulties distributing resources (availability, expertise, and support)
- techniques to measure performance indices

Why use analytical modeling in this context?

It allows to study the distributed teams’ behavior since the early phases of the project.
Globally distributed projects

- dispersed multiple sites (borders and time zones)
- difficulties distributing resources (availability, expertise, and support)
- techniques to measure performance indices

Why use analytical modeling in this context?

It allows to study the distributed teams’ behavior since the early phases of the project.
Outline

- Analytical modeling
- Team’s performance analysis
  - Single-site
  - Multi-site
- Conclusions and future works
High-level modeling formalism

- *Stochastic Automata Networks* (SAN): modular representation (states and transitions)
- applicable in many domains: Geology, Social sciences, Business, Computer science and Telecommunication

Goal

- analytical modeling of geographically distributed projects
- analysis of the interplay of interactions and productivity
Approach

Project → modeling formalism → Model → Performance indices → improve

High-level modeling formalism

- *Stochastic Automata Networks* (SAN): modular representation (states and transitions)
- applicable in many domains: Geology, Social sciences, Business, Computer science and Telecommunication

Goal

- analytical modeling of geographically distributed projects
- analysis of the interplay of interactions and productivity
Approach

High-level modeling formalism

- *Stochastic Automata Networks* (SAN): modular representation (states and transitions)
- applicable in many domains: Geology, Social sciences, Business, Computer science and Telecommunication

Goal

- analytical modeling of geographically distributed projects
- analysis of the interplay of interactions and productivity
Outline

- Analytical modeling
- Team’s performance analysis
  - Single-site
  - Multi-site
- Conclusions and future works
Single-site development team interaction pattern

Leader

Member #1

Member #2

\ldots

Member #N
Single-site development team SAN model

<table>
<thead>
<tr>
<th>Leader</th>
<th>State</th>
<th>Abstraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Mg$</td>
<td><strong>Management:</strong> managing resources, assigning responsibilities, making minor project decisions, control of tasks and schedule.</td>
</tr>
<tr>
<td></td>
<td>$Co$</td>
<td><strong>Collaboration:</strong> the leader cooperates with a member solving issues, clarifying architectural details and responding general requests.</td>
</tr>
</tbody>
</table>

![Diagram showing the interaction between $Mg$ and $Co$ with labels $a_1 \ldots a_N$ and $s_1 \ldots s_N$.]
## Single-site development team SAN model

<table>
<thead>
<tr>
<th>Member</th>
<th>State</th>
<th>Abstraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Wk$</td>
<td><strong>Working:</strong> execution of tasks, including cooperation with other members.</td>
</tr>
<tr>
<td></td>
<td>$Wt$</td>
<td><strong>Waiting:</strong> need to wait for the leader to cooperate in order to solve issues or demands.</td>
</tr>
<tr>
<td></td>
<td>$Co$</td>
<td><strong>Collaborating:</strong> member is synchronously cooperating with the leader.</td>
</tr>
<tr>
<td></td>
<td>$Rw$</td>
<td><strong>Reworking:</strong> bug corrections, code rewriting, requirements and design reviews.</td>
</tr>
</tbody>
</table>

![Diagram of the SAN model for a single-site development team](image-url)
## Single-site development team SAN model

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>me</strong></td>
<td><strong>Member’s expertise</strong>: frequency in which a member demands cooperation (level of expertise).</td>
</tr>
<tr>
<td><strong>a</strong></td>
<td><strong>Availability</strong>: leader’s availability to cooperate.</td>
</tr>
<tr>
<td><strong>s</strong></td>
<td><strong>Support</strong>: based on the time spent solving issues demanded by members (leader’s expertise).</td>
</tr>
<tr>
<td><strong>r</strong></td>
<td><strong>Rework</strong>: based on the time spent of a member reworking a task (level of experience and skills).</td>
</tr>
</tbody>
</table>

![Diagram of member interaction](Figure)
### Single-site development team SAN model

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$me$</td>
<td><strong>Member’s expertise</strong>: frequency in which a member demands cooperation (level of expertise).</td>
</tr>
<tr>
<td>$a$</td>
<td><strong>Availability</strong>: leader’s availability to cooperate.</td>
</tr>
<tr>
<td>$s$</td>
<td><strong>Support</strong>: based on the time spent solving issues demanded by members (leader’s expertise).</td>
</tr>
<tr>
<td>$r$</td>
<td><strong>Rework</strong>: based on the time spent of a member reworking a task (level of experience and skills).</td>
</tr>
</tbody>
</table>

![SAN model diagram]

- **Leader** $Mg$ interacts with **Member #1** $Co$ and **Member #N** $Co$.
- $a_1...a_N$ and $s_1...s_N$ represent various interactions and demands.
- $Wk$ represents work tasks or projects.
- $Rw$ and $Wt$ indicate rework and work time, respectively.
- $s_1(1-\pi)$ and $s_N(1-\pi)$ denote the time spent on tasks and rework, adjusted by the probability $\pi$.
- $a_1$ and $a_N$ indicate availability and support, respectively.
<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>me</td>
<td><strong>Member’s expertise</strong>: frequency in which a member demands cooperation (level of expertise).</td>
</tr>
<tr>
<td>a</td>
<td><strong>Availability</strong>: leader’s availability to cooperate.</td>
</tr>
<tr>
<td>s</td>
<td><strong>Support</strong>: based on the time spent solving issues demanded by members (leader’s expertise).</td>
</tr>
<tr>
<td>r</td>
<td><strong>Rework</strong>: based on the time spent of a member reworking a task (level of experience and skills).</td>
</tr>
</tbody>
</table>

![Diagram](attachment:image.png)
### Single-site development team SAN model

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$me$</td>
<td><strong>Member’s expertise:</strong> frequency in which a member demands cooperation (level of expertise).</td>
</tr>
<tr>
<td>$a$</td>
<td><strong>Availability:</strong> leader’s availability to cooperate.</td>
</tr>
<tr>
<td>$s$</td>
<td><strong>Support:</strong> based on the time spent solving issues demanded by members (leader’s expertise).</td>
</tr>
<tr>
<td>$r$</td>
<td><strong>Rework:</strong> based on the time spent of a member reworking a task (level of experience and skills).</td>
</tr>
</tbody>
</table>

**Diagram:**

- Leader
- $Mg$ to $Co$: $a_1 \ldots a_N$, $s_1 \ldots s_N$
- $Co$ to $Wk$:
  - $me_1$, $s_1(\pi)$, $r_1$
- $Co$ to $Wt$:
  - $a_1$, $s_1(1-\pi)$
- $Co$ to $Rw$:
  - $s_1(\pi)$, $r_1$

- Member #1
- $Wk$ to $Co$:
  - $me_1$, $s_1(\pi)$
- $Co$ to $Wt$:
  - $a_1$, $s_1(1-\pi)$
- $Co$ to $Rw$:
  - $s_1(\pi)$, $r_1$

- Member #N
- $Wk$ to $Co$:
  - $me_N$, $s_N(\pi)$
- $Co$ to $Wt$:
  - $a_N$, $s_N(1-\pi)$
- $Co$ to $Rw$:
  - $s_N(\pi)$, $r_N$
### Single-site development team SAN model

<table>
<thead>
<tr>
<th>Event</th>
<th>Estimated durations for an eight-hour workday</th>
</tr>
</thead>
</table>
| **m(e)** | *(I)nexperienced:* cooperation at each 90 minutes.  
*(E)xpert:* cooperation *once a day*. |
| **s** | *(L)ow:* leader requires 90 minutes for solving issues demanded by members.  
*(H)igh:* leader requires 30 minutes for solving issues demanded by members. |
| **a** | *(A)vailable:* leader cooperates every 30 minutes.  
*(B)usy:* leader presents low availability due to management duties (*once a day*). |
| **r** | *(R)ework:* member requires 120 minutes to review/correct its tasks. |
Single-site development team SAN model results

- Eight possible scenarios:
  - ILA, ILB, IHA, IHB
  - ELA, ELB, EHA, EHB

Example: EHA scenario

- Team with *experienced* members (E)
- Leader that provides a *high* (H) support
- Leader often *available* to cooperate with the members (A)
Single-site development team SAN model results

EHA/ELA scenarios: leader’s responsibilities analysis

- team with *experienced* members (E)
- leader that provides a *high* (H) or *low* (L) support
- leader often *available* to cooperate with the members (A)
Single-site development team analysis

Impact on the productivity considering members’ expertise and leader’s availability
Impact on the productivity considering members’ expertise and leader’s availability
Impact on the productivity considering members’ expertise and leader’s availability
Outline

- Analytical modeling
- Team’s performance analysis
  - Single-site
  - Multi-site
- Conclusions and future works
Multi-site development team interaction pattern

Central Team

Site #1

- Leader #1
- Member #1
- ... Member #N

Site #2

- Leader #2
- Member #1
- ... Member #N

Site #S

- Leader #S
- Member #1
- ... Member #N

...
## Multi-site development team SAN model

<table>
<thead>
<tr>
<th>State</th>
<th>Abstraction</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Mg</em></td>
<td>Similar to the single-site model.</td>
</tr>
<tr>
<td><em>Co</em></td>
<td></td>
</tr>
<tr>
<td><em>Ex</em></td>
<td><strong>External collaboration</strong>: cooperation between leader and central team.</td>
</tr>
</tbody>
</table>

![SAN model diagram]

- $a_1...a_N$ from $Mg$ to $Co$
- $l_{e1}...l_{eN}$ from $Co$ to $Ex$
- $s_{1}...s_{N}$ from $Ex$ to $Co$
- $es_{1}...es_{N}$ from $Co$ to $Mg$
# Multi-site development team SAN model

<table>
<thead>
<tr>
<th>State</th>
<th>Abstraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wk, Wt, Co, Rw</td>
<td>Similar to the single-site model.</td>
</tr>
<tr>
<td>Wr</td>
<td><strong>Waiting response:</strong> member is waiting for an external response in order to resume an specific work/task.</td>
</tr>
</tbody>
</table>

---

![Diagram of SAN model](image-url)
## Multi-site development team SAN model

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$me, a, s, r$</td>
<td>Similar to the single-site model.</td>
</tr>
<tr>
<td>$le$</td>
<td><strong>Leader’s expertise</strong>: level of expertise of the leader to deal with members’ requests without any external support.</td>
</tr>
<tr>
<td>$es$</td>
<td><strong>External support</strong>: central team ability (experience, time-zones and cultural diversities barriers) to solving issues demanded by leaders.</td>
</tr>
</tbody>
</table>

![Diagram of Multi-site development team SAN model]
### Multi-site development team SAN model

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$me, a, s, r$</td>
<td>Similar to the single-site model.</td>
</tr>
<tr>
<td>$le$</td>
<td><strong>Leader’s expertise:</strong> level of expertise of the leader to deal with members’ requests without any external support.</td>
</tr>
<tr>
<td>$es$</td>
<td><strong>External support:</strong> central team ability (experience, time-zones and cultural diversities barriers) to solving issues demanded by leaders.</td>
</tr>
</tbody>
</table>

---

**Leader**

1. $Mg \rightarrow Co \rightarrow Ex$
2. $s_1 \ldots s_N$
3. $a_1 \ldots a_N$
4. $l_1 \ldots l_e_N$
5. $e_s_1 \ldots e_s_N$

**Member #1**

1. $Wk \rightarrow Co \rightarrow Wt$
2. $r_1$
3. $s_1(\pi)$
4. $m_e_1$
5. $e_s_1$
6. $l_e_1$

**Member #N**

1. $Wk \rightarrow Co \rightarrow Wt$
2. $r_N$
3. $s_N(\pi)$
4. $m_e_N$
5. $e_s_N$
6. $l_e_N$
### Multi-site development team SAN model

<table>
<thead>
<tr>
<th>Event</th>
<th>Estimated durations for an eight-hour workday</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>(A)vailable: leader cooperates every 30 minutes.</td>
</tr>
<tr>
<td></td>
<td>(B)usy: leader cooperates only once a day.</td>
</tr>
<tr>
<td>le</td>
<td>(I)nexperienced: leader demands cooperation with central team <em>four times a day</em>.</td>
</tr>
<tr>
<td></td>
<td>(E)xpert: leader requires cooperation with central team <em>once a week</em>.</td>
</tr>
<tr>
<td>es</td>
<td>(L)ow: central team requires <em>one day</em> for responding issues.</td>
</tr>
<tr>
<td></td>
<td>(H)igh: central team requires <em>30 minutes</em> for responding issues.</td>
</tr>
<tr>
<td>me</td>
<td>member demands cooperation <em>twice a day</em>.</td>
</tr>
<tr>
<td>r</td>
<td>member requires <em>120 minutes</em> to review its tasks.</td>
</tr>
<tr>
<td>s</td>
<td>leader requires <em>one hour</em> for solving issues demanded by members.</td>
</tr>
</tbody>
</table>
Multi-site development team SAN model results

- Eight possible scenarios:
  - AIL, AIH, AEL, AEH
  - BIL, BIH, BEL, BEH

Example: AEH scenario

- team with an *available* leader (A)
- team with an *experienced* leader (E)
- central team provides *high* support (H)
### Multi-site development team SAN model results

#### Team members’ performance: the best productivity (AEH)

<table>
<thead>
<tr>
<th>% \ N</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wk</td>
<td>64.66</td>
<td>61.29</td>
<td>57.58</td>
<td>53.66</td>
<td>49.67</td>
<td>45.76</td>
<td>42.08</td>
<td>38.70</td>
<td>−40.15</td>
</tr>
<tr>
<td>Wt</td>
<td>12.50</td>
<td>17.07</td>
<td>22.09</td>
<td>27.39</td>
<td>32.79</td>
<td>38.08</td>
<td>43.06</td>
<td>47.63</td>
<td>+281.04</td>
</tr>
<tr>
<td>Rw</td>
<td>6.47</td>
<td>6.13</td>
<td>5.76</td>
<td>5.37</td>
<td>4.96</td>
<td>4.58</td>
<td>4.21</td>
<td>3.87</td>
<td>−40.19</td>
</tr>
<tr>
<td>Wr</td>
<td>0.20</td>
<td>0.19</td>
<td>0.17</td>
<td>0.17</td>
<td>0.16</td>
<td>0.14</td>
<td>0.13</td>
<td>0.12</td>
<td>−40.00</td>
</tr>
</tbody>
</table>

---

### AEH scenario: team members’ performance

- team with an *available* leader (A)
- team with an *experienced* leader (E)
- central team provides *high* support (H)
### Multi-site development team SAN model results

<table>
<thead>
<tr>
<th>% \ N</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wk</td>
<td>18.30</td>
<td>15.14</td>
<td>12.88</td>
<td>11.19</td>
<td>9.88</td>
<td>8.83</td>
<td>7.98</td>
<td>7.28</td>
<td>−60.22</td>
</tr>
<tr>
<td>Wt</td>
<td>56.99</td>
<td>64.41</td>
<td>69.73</td>
<td>73.70</td>
<td>76.79</td>
<td>79.24</td>
<td>81.24</td>
<td>82.89</td>
<td>+45.45</td>
</tr>
<tr>
<td>Co</td>
<td>4.58</td>
<td>3.79</td>
<td>3.22</td>
<td>2.80</td>
<td>2.47</td>
<td>2.21</td>
<td>2.00</td>
<td>1.82</td>
<td>−60.26</td>
</tr>
<tr>
<td>Rw</td>
<td>1.83</td>
<td>1.52</td>
<td>1.29</td>
<td>1.12</td>
<td>0.99</td>
<td>0.89</td>
<td>0.80</td>
<td>0.73</td>
<td>−60.11</td>
</tr>
<tr>
<td>Wr</td>
<td>18.30</td>
<td>15.14</td>
<td>12.88</td>
<td>11.19</td>
<td>9.87</td>
<td>8.83</td>
<td>7.98</td>
<td>7.28</td>
<td>−60.22</td>
</tr>
</tbody>
</table>

### BIL scenario: team members’ performance
- a *busy* team leader to cooperate with the members (B)
- an *inexperienced* team leader (I)
- a team that receives a *low* support from central team (L)
Multi-site development team SAN model results

Team members’ productivity analysis

Availability impact on the team’s productivity

- available leaders (AEL/AIH)
- busy leaders (BEL/BIH)
- sensitivity of availability on the team members’ productivity
The effect of high external support (AIH/BEH)

Which type of leader is better: an *available and inexperienced* leader or a *busy and expert* one?
The effect of high external support (AIH/BEH)

- Which type of leader is better: an available and inexperienced leader or a busy and expert one?

Multi-site development team SAN model results

Team members’ productivity analysis

![Graph showing productivity analysis](image-url)

- AIH
- BEH

<table>
<thead>
<tr>
<th>Number of team members</th>
<th>AIH</th>
<th>BEH</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>61.17</td>
<td>28.30</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>32.48</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>20.19</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The effect of low external support (AIL/BEL)

- Is it also better to have an available and inexperienced leader than a busy and expert one in low external support scenarios?
The effect of low external support (AIL/BEL)

- Is it also better to have an available and inexperienced leader than a busy and expert one in low external support scenarios?

Multi-site development team SAN model results

Team members’ productivity analysis

<table>
<thead>
<tr>
<th>Number of team members</th>
<th>AIL Productivity (%)</th>
<th>BEL Productivity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>29.60</td>
<td>27.60</td>
</tr>
<tr>
<td>3</td>
<td>22.04</td>
<td>21.70</td>
</tr>
<tr>
<td>4</td>
<td>18.65</td>
<td>18.25</td>
</tr>
<tr>
<td>5</td>
<td>15.70</td>
<td>15.35</td>
</tr>
<tr>
<td>6</td>
<td>13.85</td>
<td>13.50</td>
</tr>
<tr>
<td>7</td>
<td>12.20</td>
<td>11.95</td>
</tr>
<tr>
<td>8</td>
<td>10.70</td>
<td>10.45</td>
</tr>
<tr>
<td>9</td>
<td>9.35</td>
<td>9.10</td>
</tr>
</tbody>
</table>

The productivity decreases as the number of team members increases.
Outline

- Analytical modeling
- Team’s performance analysis
  - Single-site
  - Multi-site
- Conclusions and future works
Conclusion

- evaluate communication aspects affecting team members’ productivity
- basically three parameters play a direct influence in productivity:
  - leader’s availability
  - (local/external) support
  - expertise
- trade-off between leader’s expertise and external support impacting on team members’ productivity
- analytical modeling is useful to predict behaviors before implementing a project or process
- once validated models can provide new quantitative measures only changing model parameters
Future works

- analysis of the impact of centralized control mechanisms
to capture other important dimensions such as different
time-zones, geographic distance, communication patterns, teams
coordination
to model software development processes focusing on the flow of
requirements engineering, development and testing, project
schedule evolution
Thank you for your attention.