Follow-The-Sun Methodology
in a Stochastic Modeling Perspective

Ricardo M. Czekster, Paulo Fernandes, Rafael Prikladnicki,
Afonso Sales, Alan R. Santos, Thais Webber

Pontifícia Universidade Católica do Rio Grande do Sul
Av. Ipiranga, 6681 – Prédio 32 – 90619-900 – Porto Alegre, Brazil
{ricardo.czekster, paulo.fernandes, rafael.prikladnicki}@pucrs.br
afonso.sales@pucrs.br, alan.ricardo@acad.pucrs.br, thais.webber@pucrs.br

Abstract—Stochastic modeling of globally distributed projects has become a way to evaluate the performance of teams working in different time zones. The interest in extracting and analyzing quantitative data from geographically dispersed teams has grown throughout the years as major development companies were attracted by potential benefits. We direct our attention to Follow-The-Sun (FTS), a special case of globally distributed projects, where work is handed off every day from one development site to the next on a different time zone. The main benefit of FTS is to reduce project duration providing continuous software development. Our objective in this paper is to discuss a formal mapping of FTS characteristics to a stochastic model in order to predict performance indices of teams such as availability and risk assessment. The modeling effort aims to enhance understanding and feasibility evaluation for FTS projects.

Keywords—Global Software Development; Follow-the-Sun; Performance Evaluation; Stochastic Modeling.

I. INTRODUCTION

Companies are looking for ways to re-structure their IT area extending operations to offshore software development centers. Furthermore, there are research challenges in the software development field such as heterogeneous team configurations distributed in separate sites, or even different time zones, inter-cultural factors, different experience levels and technical background [1], [2], [3], [4].

Theoretically, with the use of 24-h software development, the software development cycle time can be significantly reduced. This kind of development may become a valid choice for software companies that can be applied in numerous projects [5].

Follow-The-Sun (FTS) is essentially focused on speed where the project is designed to reduce cycle-time (also known as time-to-market reduction, or duration reduction) [6]. FTS is a subset of global software development and shares many issues and challenges with global software development in terms of coordination, culture, and communication [6].

In some domains such as health-care and air traffic control, working throughout the day is a common practice [1]. In respect to software engineering, however, the benefits of around the clock development can be achieved through work transfer across time zones, thereby eliminating the need for unsociable working hours [1].

On this context, new methodologies are needed to evaluate the hand-off process in terms of performance and its impact on the overall project duration, considering communication issues and details of the task being shifted to the next site to be continued.

Carmel et al. [6] proposed a long-term research strategy for FTS, with the goal of improving measures. One of the action points identified was to develop research frameworks for time separation and to create mathematical and simulation models to better understand how to optimally architect FTS and how to evaluate its costs and benefits.

The objective of this paper is to present a formal mapping of FTS characteristics to a stochastic model in order to evaluate teams in FTS projects. Moreover, we demonstrate the benefits of using the Stochastic Automata Networks (SAN) formalism for the modeling and evaluation of distributed teams, since SAN provides a modular description with functional primitives. The modeling effort presented aims to enhance understanding and feasibility evaluation for FTS projects calculating probabilities for availability and project risk factor, such as hand-off efficiency, considering the conceptual framework proposed by Carmel. Conclusions point out future works directed to model extensions to capture advanced characteristics such as cultural issues, different levels of expertise and diverse communication problems.

II. FOLLOW-THE-SUN METHODOLOGY

Follow-The-Sun (FTS) development can be simply defined as software related tasks being hand-off from one site to the next, on a daily basis [7], [8], where 24-hour development can be achieved considering time differences among teams. Each team work on its own regular workday and, at the end of its shift, the task is handed off to the next team, which continues the work from the point the earlier team left off. In the context of global software engineering, Treinen and Miller-Frost [8] have discussed the time difference as an advantage to distribute the teams to achieve a 24-hour software development environment.

FTS is recently also referenced in the literature as round-the-clock [6], [9] and around-the-clock [10], [9]. The round-the-clock and around-the-clock are related to the 24-hour develop-
ment of a given project, in different work shifts. Visser [9] has defined FTS as specific software projects where development teams are spread across multiple time zones, focusing on the sharing of source code and relevant information at the end of each shift, summarizing 24 hours of continuous development.

Additionally, FTS development can be extended to different tasks besides the development. Gupta et al. [11] have introduced the 24-hour continuous development, considering the concept of knowledge being transferred, for example, among development and test teams. The tasks handed off between distributed teams, at the end of each workday, are basically related to the knowledge about the software itself and its requirements. Carmel, Dubinsky and Espinosa [6] have proposed FTS development as a kind of global knowledge workflow that can reduce project duration, depending on factors such as hand-off efficiency, effective communication and coordination.

In practice, FTS usage was firstly documented in 1997 when IBM decided to accelerate the development process in a given multi-site project [3]. The project was composed of five teams located on development centers of five distinct countries. Many problems were found during project execution, especially during the daily hand-offs, forcing the use of another GSD strategy, instead of FTS, to accomplish the tasks. But, FTS is being experimented to reduce development phase duration, opening research opportunities in the field of team building, global software development tools, and coordination strategies.

III. STOCHASTIC MODELING APPLICATION

Analytical modeling in software engineering contexts has been successfully used to provide quantitative performance measures [12], [13]. Stochastic models and simulation schemes have been developed towards the evaluation of software project dynamics [14], analysis of teams productivity variability [2], and different performance predictions about geographically dispersed teams [3], [15], [16], [17]. However, advances are still needed for the quantitative evaluation of teams using stochastic modeling as a tool [18].

Analytical modeling formalisms are commonly applied to describe systems in a state-based approach. Markov chains [19] and Markovian-based formalisms (such as, Queueing Networks [20] and Stochastic Automata Networks (SAN) [21], [22]) are already employed in several domains, such as economics, physics, engineering and bioinformatics, to cite a few. Nevertheless, SAN allows the stochastic modeling of systems in a high-level description format for representing a given system by subsystems (i.e., components). The modeled subsystems can have an independent behavior and occasional interdependencies [21], not necessarily strict to the modeling of queueing behavioral aspects. Moreover, SAN formalism presents several computational advantages related to modularization and powerful features, such as functional dependencies and synchronizing events [22].

Furthermore, SAN is specifically suitable for modeling globally distributed projects due to the fact that teams can be abstracted in a modular manner with a discrete set of states and events. In essence, a component (or module) is described by a stochastic automaton, where the transitions between states are labeled with probabilistic and timing information. A SAN model also has a set of events, which can change the state of one or more automata. Each event has an estimated rate (or duration), which indicates how often the event occurs. Numerically solving an analytical model, one can obtain its steady-state probabilities [19] and, from these probabilities extract measures of interest about the system under evaluation (e.g., team’s performance indices, resources availability, costs, and reliability). In addition, a SAN model easily allows to investigate different team configurations only changing input parameters to achieve better performance results [23].

Traditionally, scientific works on software engineering area present different applications for mathematical models, e.g., automated software testing processes [24], [25], and quantitative evaluation of software development teams also evaluating project risks [26], [27]. Global software development modeling is a challenge already discussed by Czekster et al. [18]. On this context, the effort on mapping real characteristics of FTS projects to a stochastic model requires the modeling of factors, such as time zone differences among teams (sites), reporting time on the end of a work shift and the catching up time on the start of other work shift. As occurs in any other modeling exercise, the difficult part in modeling is to identify all variables (or the most representative ones) that should be abstracted as well as figure out their occurrence rates and/or probabilities. Specifically on FTS projects, the initial model mapping should consider hand-off durations and the frequency related to other activities to perform some basic scenario evaluations. Next section presents the FTS mapping to a stochastic model description using SAN.

IV. FTS MAPPING TO STOCHASTIC MODELS

This section presents a mathematical modeling of specific aspects of the conceptual framework for FTS proposed by Carmel et al. [28]. The framework aims to help the analysis of the conditions where FTS can be beneficial in terms of reducing the project duration, i.e., time-to-market.

In this work we propose the mapping of FTS to an analytical model using the SAN formalism primitives and available tools for numerical solution [29], [30]. The strong feature of SAN is to model parallel entities where the problem can be modularized, easing the analysis of multi-site projects, where work is divided among geographically dispersed teams. Sometimes, teams must synchronize activities or assign new tasks, revisions or tests to be performed in other sites.

We have centered this first modeling effort on the hand-off efficiency concept [28] to evaluate the probabilities of each site being reworking activities or waiting for clarifications in a FTS project. Usually FTS projects have strict rules imposed on their interactions using asynchronous hand-offs between teams at fixed intervals of time.

Evidently, this attempt to model these contexts is far from being conclusive, but provides some insights about the modeling to measure the impact of hand-off efficiency on the time spent working/reworking by each site or team. For the
composition of our model’s states and events, we have made assumptions related to FTS characteristics and behavior (refer to Section II) as follows.

We assume in our abstraction that the workload (workday task) is equally complex for all distributed teams and that all the resources needed by a project are assigned at the start and readily available throughout its conclusion.

Also we are not considering for this initial research proposal how factors such as different time zones, different cultures, software quality aspects, teams size, tasks complexity and workload impact on the FTS model.

We instantiate a stochastic model having a set of $s$ multiple sites (e.g., each site or team encapsulating $n$ members) working in different time-zones. We are considering multi-variate conditions, where sites can shift their work in an efficient (or inefficient) manner, depending on the project configuration.

The modeling assumptions for the FTS formal mapping in a stochastic automata network can be described as follows:

- each site acts as an agile team entity that works in a set of tasks, reworks tasks, spends some time closing a workday in hand-off, and stays off-line until the next workday in hand-off opening;
- different time zone for each site or team (2, 3 or even 4 sites in a 24-hour development);
- homogeneous task size - related to a workday task - task time assuming 6-hour per site, for instance, in a 4-site FTS project;
- after hand-off opening in a site, there is a probability for reworking tasks, as well as after a team spend some time working, there is a probability for reworking, and a probability for hand-off closing (e.g., given team’s expertise);
- there is an online shared repository which allows sites to commit work at the end of the workday;
- hand-offs are asynchronous (the work is shifted without direct teams’ interactions, meaning the current team goes off-line after the hand-off process ended enabling to shift the work to the next team).

GSD projects have become an important research issue in IT companies [1], [2], [3], however there is still a lack of concern in relation to teams’ performance globally distributed. On this context, stochastic modeling is a feasible and flexible alternative to abstract GSD teams configuration in environments such as follow-the-sun to measure teams’ performance. Moreover, the stochastic modeling can be used during project planning phases in order to find relevant evidence in terms of performance analysis that could help the planning improvement.

The motivation for the modeling exercise is to provide a performance modeling and evaluation process by presenting an extensible mathematical model where its utilization is described in a few easy steps. The only concern for modelers will be directed towards the parametrization of the model with their own measures and how to extend it with new event compositions. Once the model is created, the next step is its solution, a mathematical procedure that converts the textual representation into numerical results.

The computed performance indices are actually a set of quantitative measures from the model, i.e., transient or steady-state probabilities that allow us to estimate the impact of the hand-off efficiency on the overall project duration. If, however, some other FTS feature needs analysis (other than hand-off efficiency, the focus of our approach in this paper), the model could be used to inspire future instantiations of different models. In this sense, the approach adopted here allows a minimal FTS setting where modelers can extend it at will to convey other behaviors or to test novel configurations (perhaps optimal ones or settings where teams begin to somehow be underperformed).

A. Asynchronous hand-off SAN model

Fig. 1 shows an example of FTS composed of three sites across the world. In the asynchronous hand-off point of view, the sites are unable to interact with the site that shifted the work due to different time zones, for example.

The stochastic automaton representing each site is a five-state model composed of states that can be assumed by the site at each time in a workday:

- **Off-line** state, the site is off-line due to time zone, so it is unreachable for collaboration, communication or interaction;
- **Opening** hand-off state, beginning of the workday;
- **Working** state, actual work is being performed in the site (active workday);
- **Reworking** state, misunderstandings occurred mainly in the hand-off process lead to rework [31];
- **Closing** hand-off state, end of the workday.

Fig. 2 shows the stochastic automata network model representing three sites in a FTS configuration. This model is composed of three sites, but it can be instantiated for 2, 3, or 4 sites (e.g., considering 24-hour software development with 6-hour shift per site) for a model with four sites. The individual behavior of a given Site $i$, where $i$ is the index of the site, is represented by the following events: $open_i$, $wk_i$, $rw_i$, $nt_i$, $cl_{wk_i}$, $cl_{rw_i}$, and $off_i$. Following we explain in detail each event dynamics for enabling transitions in the mathematical model:
• open$_i$ – local event for transferring the work/knowledge to the next site after the closing hand-off be completed. This event enables the transition from Off-line to Opening hand-off state in the Site #$_i$, meaning that the site is starting a workday;
• wk$_i$ – local event in a Site #$_i$ responsible for enabling the transition from the Opening hand-off state to the Working or Reworking states with a given probability $\pi_i$ or probability $1-\pi_i$, respectively;
• rw$_i$ – local event responsible for enabling the transition from Working to Reworking state, where a task must be reworked;
• nt$_i$ – when a rework is performed, this local event is responsible for enabling the transition from Reworking to Working state, representing a new task starting to be executed;
• cl$_{wk}$$_i$ – local event in a Site #$_i$ that enables the transition from Working to Closing hand-off state;
• cl$_{rw}$$_i$ – local event representing the change from Reworking to Closing hand-off state in a Site #$_i$;
• off$_i$ – after close a hand-off, this local event in a Site #$_i$ is responsible for enabling the transition from Closing hand-off to Off-line state, meaning that the site is closing a workday;

In this SAN model (Fig. 2), event wk has an associated function that allows (or not) the occurrence of this event. Since we are modeling an asynchronous hand-off, a site can open its hand-off process without changing the state of earlier or next sites. However, each site has a dependency of the state from the earlier site to start to work or rework a task. In this asynchronous model, a function is associated to event wk, which verifies if the earlier site’s hand-off has been completed. For instance, event wk$_1$ can occur if Site #3 is in Off-line state, i.e., if Site #3 has closed its hand-off. Otherwise, event wk$_1$ cannot occur and need to wait the closure of hand-off from Site #3.

The maximum number of sites in a daily cycle is finite, due to the 24 hours in a day. More sites provide more working capacity, however also require more overhead and increase the likelihood of mistakes [32].

With a better understanding of the benefits and constraints, adequate communication and coordination environments can be developed to support the tight coordination that is necessary to reap the benefits of the 24-hour model [33].

According to Jalote and Jain [33] a more realistic model will consider time slots that are overlapping. However in this paper we focus our attention on the asynchronous model and a synchronous model will be part of a future research.

This section presented the mapping of the interaction pattern of development sites under FTS methodology for a SAN model. The main entities (Site #1, Site #2 and Site #3) are represented by automata. The abstraction, for instance, represents the hand-off within a project: some of the time is spent on Opening hand-off, Closing hand-off and the rest in Working or Rework states.

V. SAN MODEL PARAMETERISATION

An important phase of analytical modeling is the estimation of event rates, because the dynamic of analytical modeling is given by the event transitions between model’s states.

In a SAN model, an event is the entity responsible for the transition between model states. Local type events (loc) are events that change a local state of a given automaton. Synchronizing type events (syn) are events that simultaneously change states of one or more automata and they are used to describe the interaction between two or more automata.

Table I shows examples of durations for the events in the Follow-The-Sun model previously presented, assuming as reference an eight-hour workday.

These values were chosen considering average values selected ad-hoc. The values of probabilities $\pi_1$, $\pi_2$ and $\pi_3$ are defined for example as 0.5, meaning in average that 50% of the time the site continues a task from the earlier site and 50% needs to rework a task.
TABLE I
ESTIMATED EVENT RATES - ASYNCHRONOUS HAND-OFF MODEL

<table>
<thead>
<tr>
<th>Type</th>
<th>Event</th>
<th>Description</th>
<th>Average time</th>
</tr>
</thead>
<tbody>
<tr>
<td>loc</td>
<td>open</td>
<td>After being off-line for (in average) 16 hours, a site initiates the hand-off opening process.</td>
<td>16h</td>
</tr>
<tr>
<td>loc</td>
<td>wk</td>
<td>A site spends in average 1 hour in the beginning of the workday performing the hand-off opening process.</td>
<td>1h</td>
</tr>
<tr>
<td>loc</td>
<td>rw</td>
<td>A site works in average 1 hour per workday before reworking a pending issue.</td>
<td>1h</td>
</tr>
<tr>
<td>loc</td>
<td>nt</td>
<td>A site remains reworking a task in average 0.5 hour before starting a new task.</td>
<td>0.5h</td>
</tr>
<tr>
<td>loc</td>
<td>cl-wk</td>
<td>Before starting the hand-off closing process, a site remains working in average 4 hours per workday.</td>
<td>4h</td>
</tr>
<tr>
<td>loc</td>
<td>cl-rw</td>
<td>Before going to a hand-off closing process, a site stays in average 0.5 hour reworking a task.</td>
<td>0.5h</td>
</tr>
<tr>
<td>loc</td>
<td>off</td>
<td>A site spends in average 1 hour in the end of the workday executing the hand-off closing process.</td>
<td>1h</td>
</tr>
</tbody>
</table>

Notice that some issues should be also considered for modeling states and related parameterisation:

- the time spent in actual work, or more specifically, the probability of a team being working, does not mean an index of productivity, but represent the amount of time in a workday a team has (in average) to be productive, which can be a guide for estimating project durations combined to a quantitative analysis of team’s expertise;

- if the proposed model aims to predict overlap on team work shifts (in case of synchronous hand-off), it should be adapted to allow opening hand-off activity concurrent to the closing hand-off activity on the earlier site. Another possibility is to consider that teams agreed upon being flexible in terms of extending their working shifts for tasks clarification. Literature studies [8], [9] point out cases where is recommended a sticky hand-off, i.e., intense interactions are more favorable than a clean hand-off (drop-and-go approach);

- to support team design decisions the model is easily extended to more or less sites, with different working hours, reworking probabilities, average time spent in hand-off activities, given a specific project (in this case is important to quantify team knowledge or past experiences). The modeler need to understand more deeply the project and the participants’ profile to collect significant data from interviews, surveys or historical data in companies databases;

- cultural issues are not explored quantitatively in this paper, but remark that some risky aspects such as frequent clarifications can be related to this matter, as well as related to some lack of communication or coordination support, physical distance, level of domain knowledge, to name a few.

Once we have the model with states, transitions, events and their associated rates, the next stage is to extract the model’s performance indices, running our proposed model on specialized software tools [29], [30]. The basic indices we could extract are the steady-state probabilities of each site being in the Working or Reworking states, since can occur some delay in the hand-off process. Using the probabilities of being in the Opening hand-off state we can extract costs and calculate delays related to the earlier site, as well as the probability of being in Closing hand-off state can give us indications that many clarifications are needed or, in the best case, a team is working faster on its tasks than the next site.

For this further analysis, new states could also be added in the model (expanding the current state abstraction) in order to verify other project aspects and possible bottlenecks.

Moreover, the impact of a late hand-off closing process on the working activity in a distributed site, or the probability of a hand-off opening process that lasts more than a given duration, both can be measured using our simple model, including comparisons to other hand-off configuration scenarios.

VI. CONCLUSION

In this paper we have presented an abstraction of how to model development teams in Follow-The-Sun (FTS) context and we have proposed a scenario that enables the investigation of overall FTS project performance. This is an initial approach that showed how the combination of stochastic modeling and FTS could be mixed to produce important considerations to be used by decision makers, i.e., solving the stochastic models it is possible to obtain evidence of bottlenecks in a given project before the project execution.

We are aware that this paper does not capture all important dimensions of FTS projects such as specific cultural and language barriers, communication patterns considering social networks, or other difficulties in teams coordination. Stochastic Automata Network (SAN) allows new system compositions and behaviors by simply appending new states and relationships among previously defined entities. This is the major characteristic of SAN where modelers can profit these features to easily depict more intricate behavior. The main contribution of our work is to propose a conceptual modeling for the analysis of the hand-off efficiency for projects in a FTS global software development context. As future work users can also numerically analyze the system major bottlenecks, i.e., where the lack of hand-off is bound to happen for a given project setup.

The proposed model focused on the hand-off efficiency and discussed how to parameterize asynchronous models with different measures or observations from different project aspects.

In this sense, we would like to stress that the model proposed here could be extended to a creation of new synchronous model and re-parameterize to inspire new analysis and new configurations. It could be also used, for instance, to numerically evaluate the effectiveness of FTS itself and which types of projects are more suitable for this approach and the ones where global software development would perform better.
REFERENCES