HOW TO APPLY PROCESS MINING TECHNIQUES IN SCAMPI APPRAISALS

NAME: Arthur Valle, PhD
TITLE: Academic Staff Member
ORGANIZATION: Wintec (Waikato Institute of Technology), New Zealand
• Arthur Valle, PhD.
• I founded my own consulting company in 2000: TRENDSET Consulting
• 18+ years of experience in Software Process Improvement
• Former CMMI instructor and consultant (from ISD Brazil)
• Currently, an academic staff member at Wintec (Waikato Institute of Technology) in New Zealand
• Appraisal Team Member (ATM) in several SCAMPI Classes A, B and C
• In the beginning of my PhD degree I realized that I could reduce my effort in SCAMPI appraisals by using Process Mining
PRESENTATION OUTCOMES

• Discuss some limitations of SCAMPI appraisals/method
• Understand Process Mining (PM)
• Understand “Process Mining Extension to SCAMPI”
  • a feasible, usable and useful method that reduces some limitations of the current SCAMPI method and defines which, when, where, how and why to apply Process Mining techniques in SCAMPI-based process appraisals
  • developed as an outcome of a PhD degree (Technical Report*) and based on SSD-Service System Development process area from CMMI-SVC
• Understand the verification and validation of the extended method
• Identify benefits (and limitations) of applying Process Mining techniques in SCAMPI appraisals

CONCERNS ABOUT SCAMPI

SCAMPI appraisals/method present limitations*, such as:

- SCAMPI appraisals are inherently dependent on the appraisers and the competencies of appraisers.
- SCAMPI appraisals are long, complex, expensive and resource demanding (especially regarding data collection and analysis tasks).
- SCAMPI method presents subjectivity to analyze data and to judge about the implementation of practices.
- SCAMPI method does not provide confidence regarding sample selection and its representativeness.

* according to an online survey answered by 25 SCAMPI appraisers and CMMI experts
Process Mining is a relatively young discipline that sits between data mining and process modeling & analysis.

The idea is to discover, monitor and improve real processes (i.e., not assumed processes) by extracting knowledge from event logs readily available in today’s information systems (VAN DER AALST, 2011).

It can be applied to aid collecting and analyzing data in SCAMPI appraisals.

However, there is a lack of a method that defines what, when, where, how and why to apply Process Mining techniques in process appraisals.

source: www.processmining.org
OPPORTUNITIES

• To **reduce amount of effort and time** for data collection and analysis activities in SCAMPI appraisals.

• To **assess more process instances** than are typically done in current SCAMPI appraisals (i.e. increasing sample size).

• To **reduce subjectivity** when analyzing objective evidence and judging about reference model practices implementation.

• To use Process Mining algorithms and tools **to automatically identify the actual process being performed and its deviations from organizational policies, standard processes and reference models practices.**

• To **enhance data analysis techniques** in current SCAMPI method in order to also evaluate aspects regarding who (and when) performed an activity, and in which order.
REQUIREMENTS FOR THE METHOD

Requirements:

• Extended method should reduce the dependency on appraisers and their competencies.

• Extended method should reduce the amount of effort and time of appraisal team (especially regarding data collection and analysis tasks).

• Extended method should reduce subjectivity to analyze data and to judge about the implementation of model practices.

• Extended method should increase confidence regarding sample selection and its representativeness.
“Process Mining Extension to SCAMPI” is an extension to SCAMPI\textsuperscript{SM} v1.3b: Method Definition Document for SCAMPI A, B, and C.

It is an extended method that adds **Process Mining techniques** - such as process discovery and conformance checking - into SCAMPI method to provide an explicit and focused basis for appraising an organization using such techniques **aiming to lead to more reliable and effective appraisals**.

It **does not intend to replace SCAMPI MDD**. Rather, it focuses on how Process Mining aspects affect the main content of SCAMPI, e.g. its processes and activities.
Underlined activities have extended content to current SCAMPI method.

**Bold** processes and activities are new content to current SCAMPI method.

Note that although in the chart only processes and activities are presented, any derived element (such as inputs, outputs, tools and techniques) could have been extended.
1.1 Plan and Prepare for Appraisal

1.1.1 Determine Appraisal Objectives

Activity Description
- Analyze the extent to which the process execution, as recorded in the event log, corresponds to the de Jure process model (either reflecting the organization’s standard process or the defined process). The purpose of this activity is to pinpoint deviations and quantify the level of compliance.

Required Practices
- The appraisal team member(s) acting as Process Mining analyst(s) shall do the following:
  - (ABC) load event log (if not already there)
  - (ABC) load de Jure model
  - (ABC) map activities in the event log to activities in the de Jure process models.
  - (ABC) check results
  - (ABC) calculate fitness measure
  - (ABC) identify process instances that deviate from de Jure model, and how

1.1.2 Determine Data Collection Strategy

Parameters and Limits
- (ABC) The appraisal team must evaluate the content of Process Mining artifacts to determine how they support model practice or model component implementation.
- Note that two different de Jure models could exist. One reflects organization’s standard processes while the other reflects defined processes.
- Appraisal team member(s) performing process analyst role should identify which Process Mining algorithm to apply. There is a recommendation to use Conformance Checking algorithm. For a reference about main Process Mining algorithms see appendix F.
- As a reference, value for fitness indicator should be as high as possible (i.e. a value higher than 0.8 is desired).

1.1.3 Determine Appraisal Constraints

Implementation Guidance
- Conformance can be checked by using metrics to determine the extent to which the behavior observed in the event log complies with the de Jure model. Such analysis is performed by establishing a one-to-one mapping.
PROCESS MINING EXTENSION TO SCAMPI: HOW IT WORKS

(new) Process Mining Activities:
1.A.1 - Obtain Process Mining Artifacts;
1.A.2 - Obtain Process Mining Elements;
2.A.1 - Familiarize and Filter Event Log;
2.A.2 - Discover Actual Process from Event Log; 2.A.3 - Check Conformance of Event Log with de Jure Model;
2.A.4 - Compare Conformance between de Facto model and de Jure Model;
2.A.5 - Check Conformance to Business Rules; 2.A.6 - Examine Process Mining Results.

Process Mining Elements:

*de Facto model*: process model that reflects the actual process being performed.

*de Jure model*: process model that reflects the expected behavior of a process.
The application of the extended method was done through:

- **Case Studies** (applying the method in real situations to verify the method)
- **Experts’ Review** (to validate the method by obtaining a judgment from specialists regarding the usability, feasibility and utility of the proposed method)
Conduction of 2 SCAMPI C appraisals using the extended method in 2 different organizational units of an IT company in Brazil:

**Case A** - recommended minimum number of instances: **10 process instances** (as per SCAMPI formula)

**Case B** - all instances: **All 1911 process instances available** (minimum number required by SCAMPI formula was only 4 instances).

Both cases are related to **software maintenance operations**, whose lifecycles are supported by software applications, which **record data in a format readable by Process Mining tools**.

Both SCAMPI Class C appraisals aimed to identify the adherence to selected practices of Service Delivery (SD) process area from CMMI-SVC:

- SP 3.1 - Receive and Process Service Requests
- SP 3.2 - Operate the Service System
- GP 2.1 - Establish an Organizational Policy
- GP 3.1 - Establish a Defined Process
- GP 2.7 - Identify and Involve Relevant Stakeholders
VERIFYING THE METHOD: CASE STUDY A

a) A single analyst can not perform activities "submit" and "acknowledge".
b) In a particular process instance, at least the VAPPK, QUEUED, INPROG, SUBMITTED, ACKNOWLEDGE and COMPLETE activities must be performed, in this order.
c) If there is any two link activity, a new "submit" activity should be done.
d) Service orders should follow allowed variations of the organizational standard process.
e) Every service order must be queued.
f) All service orders must be submitted, acknowledged and completed.

GP 2.1: "The organizational policy item regarding minimum conduction of certain ordered activities in each service order was not followed in some of the selected service orders".
SD 3P 3.1: "One selected service order has not queued the service request ticket".
SD 3P 3.2: "Half of the selected service orders have not conducted submit, acknowledge and complete activities in the required order".
GP 3.1: "Some selected service orders have not completely followed the defined process".
GP 3.2: "Some..."
VERIFYING THE METHOD: CASE STUDY A – PM ALGORITHMS

Case A

- Comparison between De Jure and De Facto Models
- Conformance checking between event log and De Jure Model
- Business Rules Conformance Checking
VALIDATING THE METHOD: EXPERTS’ REVIEW

An online survey answered by 50 process improvement experts.

A walkthrough demonstrating the method - how it was developed and applied – was also performed to selected respondents.

Results shown that extended method has better performance than the original SCAMPI method, as judged by the experts.

Only 10% of respondents said that they would not apply Process Mining techniques in their future appraisals.

It was possible to prove - via hypothesis testing - that the extended method:
• is feasible, usable and useful;
• reduces identified limitations of SCAMPI method.

One-Sample T: q5; q6; q7; q8; q9; q10; q11; q12
Test of μ = 3 vs not = 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SE Mean</th>
<th>95% CI</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>q5</td>
<td>50</td>
<td>3.646</td>
<td>0.694</td>
<td>0.098</td>
<td>(3.448; 3.8429)</td>
<td>6.58</td>
<td>0.000</td>
</tr>
<tr>
<td>q6</td>
<td>50</td>
<td>3.712</td>
<td>0.882</td>
<td>0.125</td>
<td>(3.461; 3.963)</td>
<td>5.71</td>
<td>0.000</td>
</tr>
<tr>
<td>q7</td>
<td>50</td>
<td>3.402</td>
<td>0.800</td>
<td>0.113</td>
<td>(3.175; 3.650)</td>
<td>3.55</td>
<td>0.001</td>
</tr>
<tr>
<td>q8</td>
<td>50</td>
<td>3.605</td>
<td>0.861</td>
<td>0.122</td>
<td>(3.360; 3.849)</td>
<td>4.96</td>
<td>0.000</td>
</tr>
<tr>
<td>q9</td>
<td>50</td>
<td>3.732</td>
<td>0.764</td>
<td>0.108</td>
<td>(3.515; 3.949)</td>
<td>6.77</td>
<td>0.000</td>
</tr>
<tr>
<td>q10</td>
<td>50</td>
<td>3.262</td>
<td>0.809</td>
<td>0.114</td>
<td>(3.032; 3.492)</td>
<td>2.29</td>
<td>0.026</td>
</tr>
<tr>
<td>q11</td>
<td>50</td>
<td>3.652</td>
<td>0.673</td>
<td>0.095</td>
<td>(3.461; 3.844)</td>
<td>6.85</td>
<td>0.000</td>
</tr>
<tr>
<td>q12</td>
<td>50</td>
<td>3.648</td>
<td>0.650</td>
<td>0.092</td>
<td>(3.463; 3.833)</td>
<td>7.05</td>
<td>0.000</td>
</tr>
</tbody>
</table>
LIMITATIONS

Limitations:

• Due to the nature of Process Mining techniques, the adequacy of the content of artifacts are not assessed through the method;

• It is not suitable to cover all process areas of CMMI models or all specific and general practices in the “covered” process areas;

• The method requires Process Mining competencies from the appraisers;

• It requires the use of specific Process Mining tools (ex: ProM, ProMimport and Disco);

• It requires that process execution data to be collected and transformed prior to the conduction phase of the appraisal;

• Performing SCAMPI appraisals aided by Process Mining is restricted to organizational units where existing information systems record process-related data in specific format, quality and content as expected by Process Mining.
CONCLUSION

Process Mining techniques can be applied with benefits in SCAMPI appraisals because they reduce limitations as:

- time and effort for data collection and analysis;
- dependence on experience and competence of appraisers;
- unreliable sampling;
- subjectivity on judgment on the implementation of CMMI practices

“Process Mining Extension to SCAMPI” positions Process Mining as a complementary approach (and substitute at some points) to the traditional SCAMPI data collection & analysis techniques but also for judging on the implementation of CMMI practices

Appraisers and process analysts can now rely on a guidance regarding identifying which Process Mining algorithms to use and how to apply them to answer the typical “questions” of software process assessment.
REFERENCES


ADDITIONAL SLIDES

- Case study B
- Future work
VERIFYING THE METHOD: CASE STUDY B

**Process related Data**

**Process Mining related activities**

**Method new activities and elements**

---

### De Facto Process Model

- **De Facto Process Model** (i.e. actual behaviour)

### De Jure Process Model

- **De Jure Process Model** (i.e. expected behaviour)

---

**Business rules**

**Event Log**

**Outcomes of PM algorithms**

**Appraisal findings**

---

<table>
<thead>
<tr>
<th>Case B</th>
</tr>
</thead>
</table>

**Process Mining related activities:**

1. A.1 Obtain Process Mining Artifacts
2. A.2 Obtain Process Mining Elements
3. A.2.1 Familiarize and Filter Event Log
4. A.2.2 Discover Actual Process from Event Log
5. A.2.3 Check Conformance of Event Log with de Jure Model
6. A.2.4 Compare Conformance between de Facto model and de Jure Model
7. A.2.5 Check Conformance to Business Rules
8. A.2.6 Examine Process Mining Results

---

**Data**

- **Event Log**

---

**Organization**

- **CMMI**

---

**Business Rules**

- **Findings**

---

**Method new activities and elements**

**VERIFYING THE METHOD:**

**CASE STUDY B**

- GP 2.1: “the organizational policy item regarding conduction of mandatory activities for each incident was not followed in almost 25% of the tickets”
- SD SP 3.1: “out of 1911 tickets have not accepted the incident after opening the incident”.
- SD SP 3.2: “none”.
- GP 3.1: “Most of all selected tickets have not completely followed the defined process”.
- GP 3.7: “In all tickets, who open the ticket also conduct other activities in the incident management lifecycle, which is not allowed”.
VERIFYING THE METHOD: CASE STUDY B – PM ALGORITHMS

Case B: Comparison between De Jure and De Facto Models

Conformance checking between event log and De Jure Model

Business Rules Conformance Checking
FUTURE WORK

Future work:

• Development of studies regarding SCAMPI method in order to enable the judgment of CMMI practices based on indicators or other quantitative criteria;

• Improvement of the extended method aiming to guide on coverage of specific practices and/or process areas of CMMI-SVC;

• Evolution of the current focus of the extended method (i.e. conformance) for a method that also examines performance and improvement aspects;

• Application of the method extended into new SCAMPI appraisal scenarios (e.g. SCAMPI B and A), to corroborate the findings presented here and allow other generalizations;

• Application of the extended method in other scenarios, such as internal and external audits;

• Automation of Process Mining techniques in SCAMPI appraisals through workflow automation tools such as RapidMiner (www.rapidminer.com).