

A new framework for defining realistic SLAs: An evidence-based approach ^{*}

Minsu Cho^{1,2}, Minseok Song², Carlos Müller³, Pablo Fernandez³, Adela del-Río-Ortega³, Manuel Resinas³, and Antonio Ruiz-Cortés³

¹ Ulsan National Institute of Science and Technology, Ulsan, Korea
mcho@unist.ac.kr

² Pohang University of Science and Technology, Pohang, Korea
mssong@postech.ac.kr

³ University of Seville, Seville, Spain
{cmuller,pablofm,adeladelrio,resinas,aruiz}@us.es

Abstract. In a changing and competitive business world, business processes are at the heart of modern organizations. In some cases, service level agreements (SLAs) are used to regulate how these business processes are provided. This is usually the case when the business process is outsourced, and some guarantees about how the outsourcing service is provided are required. Although some work has been done concerning the structure of SLAs for business processes, the definition of service level objectives (SLOs) remains a manual task performed by experts based on their previous knowledge and intuition. Therefore, an evidence-based approach that curtails humans involvement is required for the definition of realistic while challenging SLOs. This is the purpose of this paper, where performance-focused process mining, goal programming optimization techniques, and simulation techniques have been availed to implement an evidence-based framework for the definition of SLAs. Furthermore, the applicability of the proposed framework has been evaluated in a case study carried out in a hospital scenario.

Keywords: Service Level Agreement, Process Mining, Process Performance Indicators, Optimization, Goal Programming, Simulation

1 Introduction

In a changing and competitive business world, business processes are at the heart of modern organizations [1]. In some cases, service level agreements (SLAs) are used to regulate how these business processes are provided. This is usually the case when the business process is outsourced, and some guarantees about how the outsourcing service is provided are required [2].

^{*} This work was partially supported by the European Commission (FEDER), the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 645751 (RISE_BPM), the Spanish and the Andalusian R&D&I programs (grants TIN2015-70560-R, P12-TIC-1867), the National Research Foundation of Korea (No.NRF-2014K1A3A7A030737007).

Although some work has been done concerning the structure of SLAs for business processes [2], the problem of defining the actual service level objectives (SLOs), which are essential factors of an SLA denoting requirements on the service performance, in a specific business process is largely unaddressed. This issue involves first choosing the process performance indicators (PPIs) that should be considered in the SLA and, second, defining their desired target. This target must be challenging, but achievable to ensure a good process performance. A consequence of this lack of methodology for defining SLOs is that, in the current state of practice, the definition of SLOs is usually carried out by experts based on their previous knowledge and intuition, and sometimes following a trial and error model. This is far from desired, since, according to [3], definition of objectives requires a theory and a practical base and it should meet certain requirements: not being based on experts opinion, but on measurement data; respecting the statistical properties of the measure, such as measure scale and distribution, and be resilient against outlier values; and being repeatable, transparent and easy to carry out.

To overcome this problem, in this paper, we propose a framework that includes a series of steps for defining SLAs with a systematic evidence-driven approach. The proposed method covers the understanding of current behaviors of business processes, defining SLOs, deriving optimized SLOs with improvement actions, and evaluating the expected effects with a simulation. Specifically, in this paper, we present a proposal to implement the first three steps. This proposal supports a broad range of PPIs and employs performance-focused process mining, optimization techniques for multi-objective programming, and simulation techniques. The contributions of our research are as follows: (i) connecting the realms of process mining and SLAs; and (ii) proposing a new systematic approach to defining SLAs based on evidence. The applicability of our approach has been demonstrated with an experimental evaluation based on a hospital scenario.

The remainder of this paper is organized as follows. Section 2 introduces our evidence-based framework for the definition of SLAs. Then, in Section 3, we describe the first three steps in a formal way: (i) how to derive PPIs and current SLOs (CSLOs) in Section 3.1; and (ii) how to optimize SLOs in Section 3.2. Section 4 describes the effectiveness of our approach with the experimental evaluations. The summarized related works are presented in Section 5, and finally, Section 6 concludes the work and describes future directions.

2 An SLA definition framework using an evidence-based approach

In this section, we introduce the SLA definition framework using an evidence-based approach. As depicted in Fig. 1, the framework consists of 6 steps: 1) process mining analysis, 2) SLOs inference, 3) SLOs optimization, 4) simulation analysis, 5) evaluation, and 6) SLA definition. Initially, we conduct process mining analysis to calculate PPIs using event logs extracted from information

systems. On the basis of PPIs, results from process mining analysis, CSLOs are inferred. These CSLOs represent the current behavior for PPIs. After that, CSLOs are used in the SLOs optimization step that generates desired service level objectives (DSLOs) by applying some optimization techniques. The next step is to build a simulation model and conduct a simulation analysis for a scenario based on the optimization. Then, the evaluation step is performed to analyze the deviation between PPIs from the simulation and the DSLOs. In such a step, according to the evaluation result, the process can revert to either *Step 2* or *Step 3*, in the other case, it can proceed to *Step 6*. Specifically, the *SLOs inference activity* is performed in the state of a high deviation between the PPIs and DSLOs; the *SLOs optimization activity* is executed in case of a low deviation between two values. Finally, new SLAs are derived based on the DSLOs that successfully pass the evaluation step.

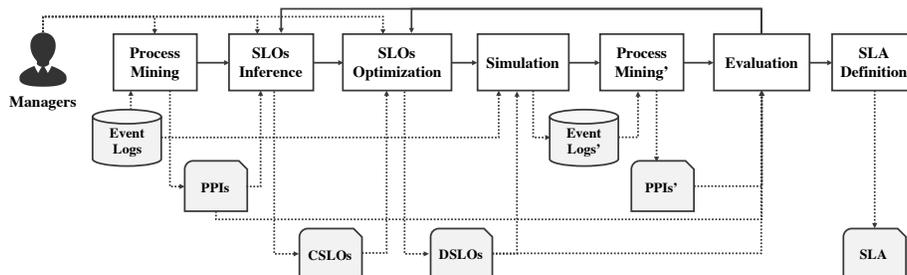


Fig. 1. Overviews of the proposed framework

As already mentioned, in this paper we focus on *Steps 1, 2, and 3* as a first approach towards supporting the whole SLA definition framework using an evidence-based approach. In order to exemplify these steps, from a user interaction perspective, Fig. 2 depicts a sequence of four mockups describing the expected interaction flow of a manager using the system in a given scenario. Specifically, the first mockup corresponds with a particular PPI selection over the outcomes of the process mining analysis (*Step 1*). Once the user selects the subset of PPIs to be optimized, the second mockup presents the essential step where current SLOs are spotted for each PPI (*Step 2*); based on the business goals, in this point, the manager can specify a desired SLO and check the appropriate potential actions to achieve the expected SLO joint with an estimated impact of the actions (as a starting point, the system calculates an estimation based on the current data that can then be tuned by the manager). Next, in the third mockup, a global set of constraints can be established typically including costs over the improvements. Finally, in the fourth mockup, the result of the optimization (*Step 3*) is shown describing the proposed improvement actions along the desired SLO and the expected metrics according to the global constraints.

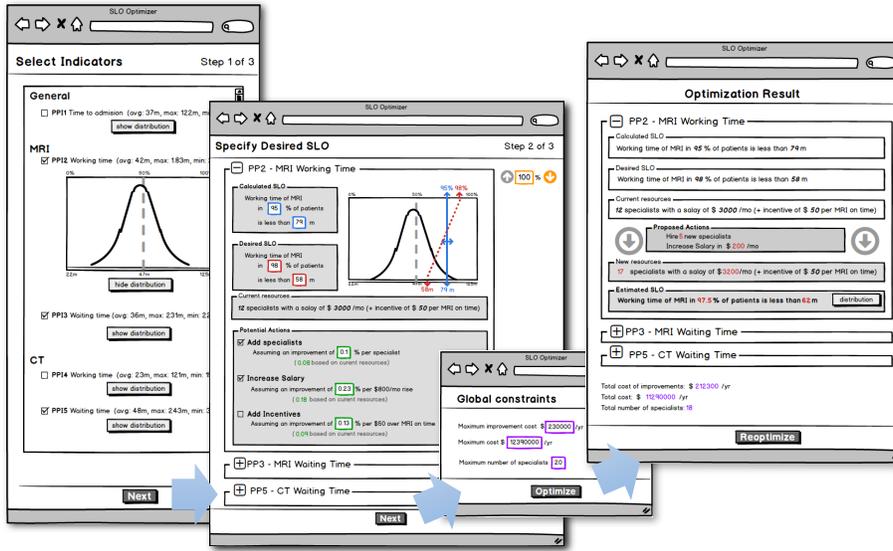


Fig. 2. User interaction flow

3 A Proposal for Obtaining DSLOs

In this section, we detail a proposal for the first three steps of the framework, which is the focus of this paper. The proposal supports a broad range of PPIs and uses a goal programming approach as the optimization technique for the SLOs optimization step.

3.1 Process Mining Analysis & SLOs Inference

Among the different perspectives involved in process mining [4], our first step focuses on the performance perspective and tries to infer the performance of a current process from its past executions stored in an event log. In such a step, a set of pre-defined PPIs (i.e., PPIs catalog) is applied, and PPIs are computed from the event log. After that, SLOs are inferred based on the calculated PPIs. In contrast to the manual approach currently followed to define SLOs, this paper proposes an evidence-based approach as an alternative. Therefore, in the SLOs inference step, managers only have a decision to select target PPIs because all PPIs are not the key performances for a process. In other words, a couple of principal PPIs are selected and inferred to CSLOs as targets to be improved. Fig. 3 provides the steps for process mining analysis and SLOs inference.

We now give a detailed explanation with formal definitions for each part. Event logs, which are the inputs of process mining, are a collection of cases, where a case is a sequence of events (describing a trace). In other words, each event belongs to a single case. Events have four properties: activity, originator,

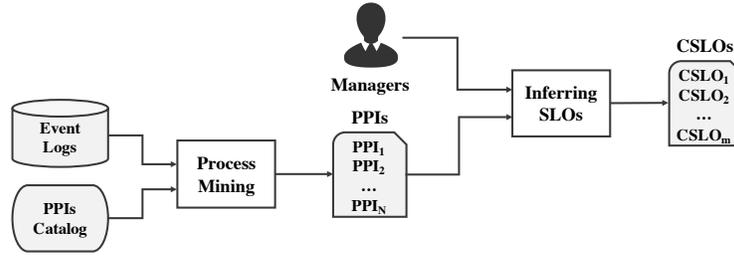


Fig. 3. Process Mining Analysis & Inferring SLOs

event type, and timestamp. Thus, events can be expressed as assigned values for these four properties. These are defined as follows.

Definition 1 (Events, Cases, and Event Log) *Let A, O, ET, T be the universe of activities, originators, event types, timestamps, respectively. Let $\mathcal{E} = A \times O \times ET \times T$ be the universe of events. Note that events are characterized by properties. For each event e and each property p , $p(e)$ denotes the value of property p for event e (e.g., $act(e)$, $type(e)$, $time(e)$ are the activity, the event type, and the timestamp of the event e , respectively). If there is no assigned value of the event e to the property p , we use \perp . Let $C = \mathcal{E}^*$ be the set of possible event sequences (i.e., cases). An event log $\mathcal{L} \in \mathbb{B}(C)$ is the set of all possible multi-sets over C .*

A simple example log is provided in Table 1. In the table, 24 events for four cases are included, and each line corresponds to a trace represented as a sequence of activities. For example, the trace of the case 1 refers to a process instance where A was started by Paul at 09:00 and completed at 10:00, B was started by Mike at 10:20 and completed at 12:00, and C was started by Allen at 13:00 and completed at 13:30. Also, event IDs are determined by the order of cases and timestamps of events (i.e., $E1: A_{Start}^{Paul,09:00}$, $E2: A_{Complete}^{Paul,10:00}$, ..., and $E24: D_{Complete}^{Allen,17:30}$). This log will be used as a running example.

Table 1. Running example log

Case	Trace
Case 1	$\langle A_{Start}^{Paul,09:00}, A_{Complete}^{Paul,10:00}, B_{Start}^{Mike,10:20}, B_{Complete}^{Mike,12:00}, D_{Start}^{Allen,13:00}, D_{Complete}^{Allen,13:30} \rangle$
Case 2	$\langle A_{Start}^{Paul,10:30}, A_{Complete}^{Paul,11:00}, C_{Start}^{Chris,12:10}, C_{Complete}^{Chris,13:00}, D_{Start}^{Allen,14:00}, D_{Complete}^{Allen,15:00} \rangle$
Case 3	$\langle A_{Start}^{Paul,12:00}, A_{Complete}^{Paul,12:30}, B_{Start}^{Mike,14:00}, B_{Complete}^{Mike,15:00}, D_{Start}^{Allen,15:30}, D_{Complete}^{Allen,16:30} \rangle$
Case 4	$\langle A_{Start}^{Paul,13:00}, A_{Complete}^{Paul,14:00}, C_{Start}^{Chris,14:30}, C_{Complete}^{Chris,15:30}, D_{Start}^{Allen,16:00}, D_{Complete}^{Allen,17:30} \rangle$

Based on the event log, we identify events to be used for calculating PPIs through two elements: entity types and entity identifiers. The entity type includes

activity and originator. The entity identifiers signify the possible values that belong to the entity type. For example, in Table 1, A, B, C, and D are the entity identifiers of the entity type activity. Based on these two elements and the log (i.e., the event log, the entity type, and the entity identifier), required events are filtered and extracted through the ψ function. After that, extracted events are calculated based on measures such as count, working time, and waiting time. A PPI ($P_n(M(\mathcal{E}))$) is defined as calculating n-th percentile (P_n) from computed measure values for the filtered events. The PPI is defined as follows.

Definition 2 (Process Performance Indicators) *Let T and V be the universe of entity types and universe of possible values, respectively. For each entity type $t \in T$, V_t denotes the set of possible values, i.e., the set of entity identifiers of type t . Let $\psi \in L \times T \times V_T \Rightarrow \mathcal{E}$ is a function that finds out the set of events from an event log for a given entity type and an entity identifier (where, \mathcal{E} is the set of events). M is the measures such as count, working time, waiting time, duration, etc. $P_n(M(\mathcal{E}))$ is the process performance indicator from an event log for a given entity type and an entity identifier, and a measure (where, $P_n =$ be the n-th percentile function). Note that P_{25} , P_{50} , P_{75} are 1st quantile, median, 3rd quantile, respectively.*

For example, we can get following examples from the Table 1;
 $\psi(L, \text{Activity}, A) = \{E1, E2, E7, E8, E13, E14, E19, E20\}$, $\psi(L, \text{Originator}, \text{Allen}) = \{E5, E6, E11, E12, E17, E18, E23, E24\}$. As an example of PPIs, the median of working time for $\psi(L, \text{Activity}, A)$ is calculated as 30 minutes from $\{60, 30, 30, 30\}$, and it also can be denoted as follows: *median of working time of the A is 30 minutes.*

The next step is to infer SLOs based on the calculated PPIs using process mining. SLO is defined as follows.

Definition 3 (Service Level Objectives and Inferring function) *Let M be the universe of measurements. x is the target value of the measurement m , and $P(t)$ is the function deriving probability of t . A SLO $P(m \leq x) \geq n\%$ is the probability(m) that measurement is less than x must be at least $n\%$. Let $I \in \Gamma(P_x(M(\mathcal{E}))) \Rightarrow \{P(m \leq x) \geq n\%\}$ be a function that infers the SLO from the PPI.*

SLO is defined as a probability that a measure of cases that have the entity identifier \leq value must be more than $n\%$. CSLOs are automatically inferred from PPIs using the I function. Overall structures of CSLOs and PPIs are quite similar; thus, we can easily establish CSLOs using given PPIs. As we explained earlier, PPI is defined as n-th percentile of a measure of an entity identifier is value. Based on the PPI, CSLO becomes a measure of an entity identifier must be less than the value in $n\%$ of cases. For example, in Table 1, one of the PPIs, the median (50th percentile) of working time of the activity A is 30 minutes (i.e. PPI_1). Then, the related CSLO becomes the working time of the activity A must be less than 30 minutes in 50% of the cases. Also, there is another PPI that the median of working time of the originator Allen is 60 minutes (i.e., PPI_2).

Then, the corresponding CSLO becomes the *working time* performed by *Allen* must be less than *60 minutes* in *50%* of the cases.

3.2 SLOs Optimization

The objective for the optimization step is to maximize the whole effect by minimizing the target value of each calculated SLO while maintaining it achievable and realistic by selecting the best improvement actions that enhance the process performance. Therefore, it needs a multi-objective programming approach to accomplish multiple goals. We employ the goal programming (GP) approach [5]. The goal programming method is one of the popular approaches for the multi-objective programming problem [5]. Fig. 4 shows the SLOs optimization step. In our approach, the inputs of the GP model are *improvement actions*, *CSLOs*, and *business constraints*. We assume that improvement actions are given based on prior knowledge or qualitative research (e.g., interviews and surveys). Employing more resources and providing incentives are a part of the typical examples of the actions. As explained in Section 3.1, CSLOs are derived from event logs. Finally, a manager has to determine demands and constraints including costs of implementation actions, expected SLOs and importance of each SLO. Here, the expected SLOs signify manager's expectation regarding the derived SLOs. On the basis of three inputs, a GP model is constructed, and the output of the model are how many and what improvement actions are used for each goal and the minimized SLOs (DSLOs).

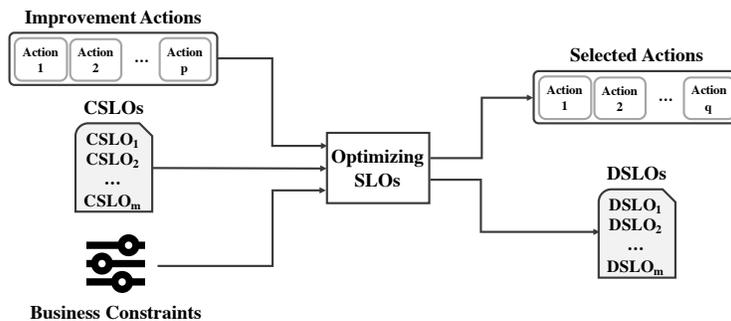


Fig. 4. Optimizing SLOs

Before explaining the GP model, we introduce the symbols that are described in Table 2.

$P_n(\mu, \sigma^2)$ denotes the percentile function for a normal distribution with mean (μ) and variance (σ^2). In general, the percentile function is defined as the infimum function of the cumulative distribution function [6]. Here, based on two aspects, we consider that percentiles are represented by a normal curve plot and

Table 2. Optimization Symbols

<i>Symbol</i>	<i>Meaning</i>
V	Number of entity identifiers in all CSLOs
M	Number of available improvement actions
i	Indices of entity identifiers, ($i = 1, 2, \dots, V$)
j	Indices of available actions, ($j = 1, 2, \dots, M$)
m	Types of measures, ($m = \{d : duration, wo : working, wa : waiting\}$)
$x_{i,j}$	Number of applications of action j for entity identifier i
$l_{i,j}$	Lower bound of number of applications of action j for entity identifier i
$u_{i,j}$	Upper bound of number of applications of action j for entity identifier i
μ_i^m	Current mean of measure m for entity identifier i
σ_i^m	Current standard deviation of measure m for entity identifier i
$f_{i,j}^m$	Effect on mean of measure m of action j for entity identifier i
$h_{i,j}^m$	Effect on std. dev. of measure m of action j for entity identifier i
$c_{i,j}$	Unit cost of method j for entity identifier i
C	Planned implement action cost
X	Target percentage by manager ($0 \leq X \leq 1$)
T	Target value by manager
W	Determined range weight for target value ($0 \leq W \leq 1$)
w_k	Importance of SLO k ($k = 1, 2, \dots, K$)
$P_n(\mu, \sigma^2)$	n-th Percentile function with μ and σ^2

can be expressed with two variables μ and σ^2 . First, there is a principle that large populations follow a normal distribution [7]. Second, the improvement actions in this paper have an effect on decreasing mean and standard deviation of distributions. Fig. 5 provides the graphical explanation. In a current distribution for an SLO, the target value based on 95% is V_1 . If an improvement action makes the mean decrease without any other changes, the distribution moves to the left. As such, the reduced new target value (V_2) is derived as provided in the left graph of Fig. 5. On the other hand, if an improvement action affects the decrease of the standard deviation, the distribution becomes more centralized than before, and the new target value (V_3) is derived as shown in the middle of Fig. 5. Furthermore, an improvement action can affect to reduction of both mean and standard deviation. Then, as shown in the right of Fig. 5, the target value is decreased as V_5 depending on the decrease of mean and standard deviation. The following is the formalization of the percentile function with the normal distribution.

Definition 4 (Percentile Function) *Let f be the probability density function, the cumulative distribution function F as follows:*

$$F(x) = \int_{-\infty}^x f(t)dt \text{ (where, } -\infty \leq x \leq \infty \text{)}$$

With reference to the function F , percentile function is

$$P(p) = \inf\{x \in \mathbb{R} : p \leq F(x)\} \text{ (where, } \inf = \text{infimum function)}$$

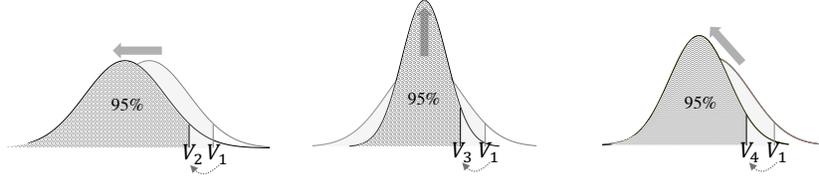


Fig. 5. Effects on target values for SLOs based on improvement actions

for a probability $0 \leq p \leq 1$. Based on the principle that large populations follow a normal distribution, percentile function becomes the inverse function of the cumulative normal distribution function. The cumulative distribution function for normal distribution with μ and σ^2 is as follows.

$$F_x(\mu, \sigma^2) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sigma \sqrt{2}} \right) \right] \quad (\text{where, erf} = \text{error function})$$

Let n -th percentile function for normal distribution with μ and σ^2 be defined as follows.

$$P_n(\mu, \sigma^2) = F_x^{-1}(\mu, \sigma^2) = \mu + \sigma \sqrt{2} \operatorname{erf}^{-1}(2p - 1) \quad (n\% = p)$$

As we explained earlier, the GP model aims at minimizing the target values of all SLOs by employing improvement actions. Therefore, an individual optimization model for each SLO is constructed. Then the GP model is formulated by combining all optimization models together. An optimization model for each goal is formalized as follows.

Definition 5 (Optimization Model for each goal)

O.F.

$$DSLO_i = \min P_X(\mu_i^{m'}, \sigma_i^{m'^2})$$

where,

$$\begin{aligned} \mu_i^{m'} &= \mu_i^m + \sum_{j=1}^M x_{i,j} f_{i,j}^m \\ \sigma_i^{m'} &= \sigma_i^m + \sum_{j=1}^M x_{i,j} h_{i,j}^m \end{aligned}$$

Constraints

$$T \times (1 - W) \leq DSLO_i \leq P_X(\mu_i^m, \sigma_i^{m2})$$

$$0 \leq x_{i,1}, x_{i,2}, \dots, x_{i,M}$$

$$x_{i,1}, x_{i,2}, \dots, x_{i,M} = \text{integer}$$

$$l_{i,j} \leq x_{i,j} \leq u_{i,j} \quad (\text{for } j = 1, 2, \dots, M)$$

$$\sum_{i=1}^V \sum_{j=1}^M x_{i,j} c_{i,j} \leq C$$

As we explained earlier, improvement actions can influence the mean and standard deviation of the distribution for SLOs. As such, the objective function is formalized aiming to minimize the percentile function considering the modified mean and the standard deviation ($P_X(\mu_i^{m'}, \sigma_i^{m'^2})$). Here, the updated mean ($\mu_i^{m'}$) and standard deviation ($\sigma_i^{m'}$) are described as the difference between the current values (μ_i^m and σ_i^{m2}) and the effects of the applying improvement actions

(i.e., $\sum_{j=1}^M x_{i,j} f_{i,j}^m$ and $\sum_{j=1}^M x_{i,j} h_{i,j}^m$ that denote the reduction of mean and standard deviation, respectively).

For the constraints in the optimization model, the expected SLO determined by managers is included as a target value with a specific target percentage. Considering the pre-determined expected SLOs, we set the range of $DSLO_i$ that it should be less than or equal to the current value ($P_X(\mu_i^m, \sigma_i^{m2})$) and greater than or equal to the value from the target value (T) and range weight (W). Moreover, another constraint is that the number of applications for each action ($x_{i,j}$) should be bigger than 0 and integer. In this regard, we can also determine a lower bound ($l_{i,j}$) and an upper bound ($u_{i,j}$) of the number of applications for each action. Furthermore, the cost-related constraint is also included so that total used cost for implementation ($\sum_{i=1}^V \sum_{j=1}^M x_{i,j} c_{i,j}$) is less than the planned implement action cost (C).

At last, we describe how to formalize the GP model that combines the optimization model for the selected SLOs. The objective function of the GP model considers both the changes of SLOs (i.e., the difference between CSLOs and the minimized SLOs ($DSLO_s$)) and the importance of each goal determined by a manager. Also, constraints and bounds in optimization models for goals are included. Formalization for the GP model is as follows.

Definition 6 (GP Model)

$$\begin{aligned} \text{O.F.} \quad \max Z = & w_1 \frac{CSLO_1 - DSLO_1}{CSLO_1} + w_2 \frac{CSLO_2 - DSLO_2}{CSLO_2} + \dots \\ & + w_K \frac{CSLO_K - DSLO_K}{CSLO_K} \end{aligned}$$

subject to *Constraints and bounds in optimization models for goals*

4 Experimental Evaluation

To demonstrate the effectiveness of our proposed approach, we apply it to an examination process in an outpatient clinic and the corresponding log utilized in [8]. In Section 4.1, we introduce the examination process and the corresponding log applied in the evaluation. In Section 4.2, we describe the results of PPIs calculation and CSLOs conversion. Section 4.3 introduces the setup for the optimization, while Section 4.4 provides the results of optimization, i.e., DSLOs.

4.1 Experiment Design and Data set

As we introduced earlier, we used the examination flows in the outpatient clinic and the corresponding event log. Fig. 6 provides the graphical description of the examination process. In the process, patients (i.e., cases) firstly visit a hospital and get both the lab test and the X-ray test. Then, if needed, patients get the electrocardiogram test (ECG). After that, they visit the hospital again and get either computerized tomography (CT) or magnetic resonance imaging (MRI) according to the results of the tests in the first visit. Lastly, the process is finished

with the third visit of the patients. The proposed framework was applied to the corresponding log of the examination process. The log included 7000 events performed by 17 resources for 1000 cases.

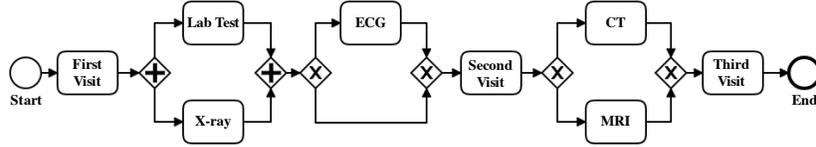


Fig. 6. The examination process used in the evaluation

In the case study, we focused on PPIs defined for the working and waiting time of the test-related activities included in the process. Also, for each indicator, we applied various aggregation functions such as median, first quartile (*1st Q*), third quartile (*3rd Q*), five percentiles (*5%*), and 95 percentiles (*95%*) to understand the distribution of the indicator. We computed PPIs with the examination event log, and Table 3 provides the results in detail.

Table 3. Calculated results of PPIs

		(measure: min.)				
Time value	Activity	Median	1st Q.	3rd Q.	5%	95%
Working Time	X-ray	20.0	19.0	21.0	17.0	23.0
	Lab Test	20.0	19.0	21.0	17.0	23.0
	ECG	30.0	27.0	33.0	22.0	38.0
	MRI	61.0	56.0	64.3	50.0	71.0
	CT	45.0	44.0	46.0	42.0	48.0
Waiting Time	X-ray	30.0	27.0	33.0	22.0	38.0
	Lab Test	30.0	26.7	33.0	22.0	38.0
	ECG	0.0	0.0	0.0	0.0	1.0
	MRI	7223.5	6931.8	7547.2	6478.5	7975.7
	CT	4314.5	3994.7	4651.2	3506.9	5089.0

4.2 Results for PPIs and CSLOs

As described in the Table 3, we identified that MRI had higher working time than any other activities (e.g., the median of working time of MRI was 61 minutes). With regard to the waiting time, a couple of activities had higher values than others: MRI and CT. These results were used to determine the candidates for optimization (i.e., CSLOs).

To decide what PPIs are taken into account for the CSLOs extraction, we can consider two types of criteria. First, the indicators that are linked to a critical

part in a process, e.g., a primary activity or a sub-process can be selected because they are necessary to improve the process. However, this approach has to be determined by a manager of an organization. In other words, it is required to have a domain knowledge of the process. The other approach is to select problematic indicators that have a high potential to be improved such as indicators that have high volatility or unexpectedly low values. Since that information from the manager was not available in our case study, we selected the second option. Among several PPIs, we selected three of them, and the corresponding CSLOs were obtained as follows.

- PPI_1 : 95th percentile (i.e., 95%) of working time of MRI is 71.0 min.
- $CSLO_1$: Working time of MRI must be less than **71.0** min. in **95%** of patients.
- PPI_2 : Median (i.e., 50th percentile) of waiting time of MRI is 7223.5 min.
- $CSLO_2$: Waiting time of MRI must be less than **7223.5** min. in **50%** of patients.
- PPI_3 : Median of waiting time of CT is 4314.5 min.
- $CSLO_3$: Waiting time of CT must be less than **43145** min. in **50%** of patients.

4.3 Setup for optimization

Based on the calculated CSLOs, we built a GP optimization model for two activities ($i = \{1 : MRI, 2 : CT\}$) and a couple of time measures ($m = \{wo : working, wa : waiting\}$) in this case study. As the inputs for the GP model, we first used the target values of CSLOs that were derived in Section 4.2: $CSLO_1 = 71.0$, $CSLO_2 = 7223.5$, and $CSLO_3 = 4314.5$. Second, we employed three improvement actions ($j=1, 2, 3$): employing more resources (*Action 1*), changing resources into more qualified people (*Action 2*), and employing managers (*Action 3*). As we explained earlier, each action has an effect on decreasing the mean and the standard deviation of time values for entity identifiers (i.e. activities in the case study). Among three actions, the action 1 lowers the average of waiting time for activities, while the action 2 reduces the mean of working time and waiting time. On the other hand, the action 3 decreases the standard deviation of working and average of waiting time. In this model, detailed effects and costs of each action are provided in Table 4.

Table 4. Effects and unit costs of each action for MRI and CT

Action	Cost	Effects on MRI				Effects on CT			
		f^{wo}	h^{wo}	f^{wa}	h^{wa}	f^{wo}	h^{wo}	f^{wa}	h^{wa}
1	1600	-	-	-1187m.	-	-	-	-1187m.	-
2	400	-1%	-	-47.62m.	-	-1%	-	-62.63m.	-
				/ -1m. of f^{wo}				/ -1m. of f^{wo}	
3	550	-	-10%	-10m.	-	-	-10%	-10m.	-
				/ -1% of h^{wo}				/ -1% of h^{wo}	

In the table, costs and effects on working time were assumed, while effects on waiting time were calculated from data. The effects on waiting time in action

1 were inferred from the M/M/c model of the queuing theory. With regard to the action 2 and 3, we calculated the reduction of waiting time according to the change in working time.

Lastly, several assumptions were encoded in the model as manager's decisions and business constraints: expected SLOs, bounds for the number of applications for each action, planned implement cost, and importance for each goal. Expected SLOs (i.e., manager's target SLOs) were assumed as follows. These values were applied as constraints in the model with the determined range weight ($W = 0.05$).

- $ESLO_1$: Working time of *MRI* must be less than **69.0** min. in **95%** of patients.
- $ESLO_2$: Waiting time of *MRI* must be less than **7000.0** min. in **50%** of patients.
- $ESLO_3$: Waiting time of *CT* must be less than **3200.0** min. in **50%** of patients.

Also, based on the current status of resources, the number of employing resources ($x_{i,1}$) and changing resources into more qualified people ($x_{i,2}$) for each activity were limited as 1 and 3, respectively. Moreover, we assumed that the planned implement cost was 3000 and the importances for all goals were the same as 0.5.

Based on these inputs, we built a GP model. The complete formulation of each goal and the GP model are presented in Table 5.

Table 5. The GP model for optimization

	Goal 1	Goal 2
O.F.	$DSLO_1 = \min P_{95}(\mu_1^{wo'}, \sigma_1^{wo'^2})$	$DSLO_2 = \min P_{50}(\mu_1^{wa'}, \sigma_1^{wa'^2})$
	$\mu_1^{wo'} = \mu_1^{wo} + \sum_{j=1}^3 x_{1,j} f_{1,j}^{wo}$	$\mu_1^{wa'} = \mu_1^{wa} + \sum_{j=1}^3 x_{1,j} f_{1,j}^{wa}$
	$\sigma_1^{wo'} = \sigma_1^{wo} + \sum_{j=1}^3 x_{1,j} h_{1,j}^{wo}$	$\sigma_1^{wa'} = \sigma_1^{wa} + \sum_{j=1}^3 x_{1,j} h_{1,j}^{wa}$
	Goal 3	
O.F.	$DSLO_3 = \min P_{50}(\mu_2^{wa'}, \sigma_2^{wa'^2})$	
	$\mu_2^{wa'} = \mu_2^{wa} + \sum_{j=1}^3 x_{2,j} f_{2,j}^{wa}$	
	$\sigma_2^{wa'} = \sigma_2^{wa} + \sum_{j=1}^3 x_{2,j} h_{2,j}^{wa}$	
	GP Model	
O.F.	$max Z = 0.5 \frac{71.0 - DSLO_1}{71.0} + 0.5 \frac{7223.5 - DSLO_2}{7223.5} + 0.5 \frac{4314.5 - DSLO_3}{4314.5}$	
subject to	$69.0 \times (1 - 0.05) \leq DSLO_1 \leq 71.0$	
	$7000.0 \times (1 - 0.05) \leq DSLO_2 \leq 7223.5$	
	$3200.0 \times (1 - 0.05) \leq DSLO_3 \leq 4314.5$	
	$0 \leq \sum_{i=1}^2 \sum_{j=1}^3 x_{i,j}$	$\sum_{i=1}^2 \sum_{j=1}^3 x_{i,j} = integer$
	$\sum_{i=1}^2 x_{i,1} \leq 1$	$\sum_{i=1}^2 x_{i,2} \leq 3$
	$\sum_{i=1}^2 \sum_{j=1}^3 x_{i,j} c_{i,j} \leq 3000$	

4.4 Optimization Results

Based on the constructed GP model, we obtained the optimal solution. Table 6 provides the optimization results for the case study. The results of optimization with the GP model recommended changing two resources into more qualified people (Action 2) and employing a manager (Action 3) for MRI. Moreover, for CT activity, employing one more resource (Action 1) was suggested. As such, the total used implement cost was turned to 2950. Also, through the optimization, all SLOs were improved. For example, the target value of $CSLO_1$ went from 71.0 minutes to 68.3 minutes. Likewise, the target values of $CSLO_2$ and $CSLO_3$ were decreased by 141.7 and 1181.1 minutes, respectively. Lastly, as a result of the combination of importance for each goal, there was a 16.7% reduction.

Table 6. Optimization results for the case study

<i>Applied improvement actions</i>
Changing resources into more qualified people (Action 2) for MRI: 2 (times)
Employing managers (Action 3) for MRI: 1
Employing more resources (Action 1) for CT: 1
<i>Total used cost</i>
2950 ($= 400 \times 2 + 550 \times 1 + 1600 \times 1$)
<i>Derived SLOs</i>
$DSLO_1$: Working time of <i>MRI</i> must be less than 68.3 min. in 95% of patients
$DSLO_2$: Waiting time of <i>MRI</i> must be less than 7076.4 min. in 50% of patients
$DSLO_3$: Waiting time of <i>CT</i> must be less than 3133.4 min. in 50% of patients

The result provided the optimal solutions in the given limited cost. In other words, it suggested the best answers for solving the problem that the current process has. Therefore, managers can acquire the direct improvement effects by applying the recommended actions into the activities in the process.

5 Related Work

Numerous research efforts have focused on proposing models for SLA definition in computational and non-computational domains [9, 10, 2], however, none of them deals with the definition of challenging while achievable SLOs. Some work has been carried out in this direction in the context of computational services. [11] proposes a methodology to calculate SLO thresholds to sign IT services SLAs according to service function cost from a business perspective, but it is useful only for SLAs that apply to the software infrastructure that supports business processes and not for business processes offered as a service. [12] describes a categorization of IT services and outlines a mechanism to obtain efficient SLOs

for them. However, they do that at a conceptual level and do not detail how they can be formalized to enable their automated computation.

Regarding the definition of data-based target values or thresholds for PPIs, [13] presents an approach to determine PPI thresholds based on the relationship of different PPIs and their values computed from the process execution data. In this approach, though, a proven relationship between certain PPIs is required in order to extract their thresholds.

Concerning our SLO optimization proposal, some related works exist in the context of process measurement and improvement. A series of proposals exist, e.g. [4, 14, 15], that identify correlations between PPIs that, eventually, can lead to the definition of process improvement actions. Also related to this is the business process redesign area, which tackles the radical change of a process to enhance its performance dramatically. In this area, a number of works have been presented where heuristic-based BPR frameworks, methodologies, and best practices have been proposed [16, 17]. The main drawback of these works concerning our motivating problem is that they are not SLA-aware and leave out of their scope the establishment of target values for the performance measures, or SLOs in the context of business processes offered as services.

6 Conclusion

This paper proposes a structured framework to define realistic SLAs with a systematic evidence-driven approach. The evaluation results obtained from its application to an examination process in the outpatient clinic have shown its applicability and the improvements on the performance of that process.

Our work has a couple of limitations and challenges. The case study adopted for validation covered only the time-related measures. Therefore, a more comprehensive approach that handles various indicators such as frequency and quality is required. Also, with regard to improvement actions in the optimization part, we applied assumptions about the types of actions, costs, and effects. As future work, we will establish more systematic improvement actions by exploring existing works and conducting interviews. In addition, we used the normal distribution-based percentile function with the normality principle. However, if we use the distribution itself (e.g., histogram), we can apply more improvement actions that modify skewness or kurtosis. Therefore, we need to develop a method to support this idea and be able to formulate those improvement actions.

Furthermore, at the beginning, we claimed that our approach aims at reducing the human involvement in the specification of SLOs, but we still need the experts for some steps to gather relevant information. Therefore, we plan to improve our approach by minimizing the human involvement as much as possible and increasing the portion of the data analysis. Finally, in this paper, we focused on the first three steps of the proposed framework. We are already working on implementing the remaining steps and a tool that supports the whole structure. Also, more case studies with real data in different contexts will be performed for further validations.

References

1. Harmon, P.: The scope and evolution of business process management. In: Handbook on Business Process Management. Volume 1.
2. del Río-Ortega, A., Gutiérrez, A.M., Durán, A., Resinas, m., Ruiz-Cortés, A.: Modelling service level agreements for business process outsourcing services. In: International Conference on Advanced Information Systems Engineering, Springer (2015) 485–500
3. Alves, T.L., Ypma, C., Visser, J.: Deriving metric thresholds from benchmark data. In: 26th IEEE International Conference on Software Maintenance (ICSM 2010). (2010) 1–10
4. de Leoni, M., van der Aalst, W.M.P., Dees, M.: A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs. *Inf. Syst.* **56** (2016) 235–257
5. Aouni, B., Kettani, O.: Goal programming model: A glorious history and a promising future. *European Journal of Operational Research* **133** (2001) 225–231
6. Wichura, M.J.: Algorithm as 241: The percentage points of the normal distribution. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **37**(3) (1988) 477–484
7. Whitley, E., Ball, J.: Statistics review 2: Samples and populations. *Critical Care* **6**(2) (2002) 143
8. Rozinat, A., Mans, R., Song, M., van der Aalst, W.M.P.: Discovering simulation models. *Information Systems* **34**(3) (2009) 305–327
9. Cardoso, J., Barros, A., May, N., Kylau, U.: Towards a unified service description language for the internet of services: Requirements and first developments. In: Services Computing (SCC), 2010 IEEE International Conference on. (July 2010) 602–609
10. Wieder, P., Butler, J., Theilmann, W., Yahyapour, R., eds.: Service Level Agreements for Cloud Computing. Volume 2506. Springer (2011)
11. Sauvé, J., Marques, F., Moura, A., Sampaio, M.: SLA design from a business perspective. *Ambient Networks* (2005) 72–83
12. Kieninger, A., Baltadzhiev, D., Schmitz, B., Satzger, G.: Towards Service Level Engineering for IT Services: Defining IT Services from a Line of Business Perspective. 2011 Annual SRII Global Conference (March 2011) 759–766
13. del Río-Ortega, A., Garcí, F., Resinas, M., Weber, E., Ruiz, F., Ruiz-Cortés, A.: Enriching decision making with data-based thresholds of process-related kpis. In: International Conference on Advanced Information Systems Engineering, Springer (2017) TBD
14. Rodriguez, R.R., Saiz, J.J.A., Bas, A.O.: Quantitative relationships between key performance indicators for supporting decision-making processes. *Computers in Industry* **60**(2) (2009) 104–113
15. Diamantini, C., Genga, L., Potena, D., Storti, E.: Collaborative building of an ontology of key performance indicators. In: On the Move to Meaningful Internet Systems: OTM 2014 Conferences. (2014) 148–165
16. Mansar, S.L., Reijers, H.A.: Best practices in business process redesign: validation of a redesign framework. *Computers in Industry* **56**(5) (2005) 457–471
17. Watson, H.J., Wixom, B.H.: The current state of business intelligence. *Computer* **40**(9) (2007) 96–99