Optimization-based Business Process Model Matching

Merih Seran Uysal[1][https://orcid.org/0000-0003-1115-6601], Dominik Hüser[1], and Wil M.P. van der Aalst[1][https://orcid.org/0000-0002-0955-6940]

[1]Process and Data Science Chair, RWTH Aachen University, Aachen, Germany

{uysal,wvdaalst}@pads.rwth-aachen.de   dominik.hueser@rwth-aachen.de

Abstract. The rapid increase in generation of business process models in the industry has raised the demand on the development of process model matching approaches. In this paper, we introduce a novel optimization-based business process model matching approach which can flexibly incorporate both the behavioral and label information of processes for the identifi-cation of correspondences between activities. Given two business process models, we achieve our goal by defining an integer linear program which maximizes the label similarities among process activities and the behavioral similarity between the process models. Our approach en-ables the user to determine the importance of the local label-based similarities and the global behavioral similarity of the models by offering the utilization of a predefined weighting param-eter, allowing for flexibility. Moreover, extensive experimental evaluation performed on three real-world datasets points out the high accuracy of our proposal, outperforming the state of the art.

Keywords: Process Model Matching, Optimization Problem, Integer Linear Programming, Be-havioral Similarity

1 Introduction

The ubiquity of advanced capabilities of the digital world enables organizations to generate and store process models which exhibit indispensable activities of their business processes in various domains, e.g., finance, logistics, and production [1, 14, 18]. The resulting increase in uptake of business process model repositories leads to the need for the development of techniques in various fields, e.g. storage of process models, management of repositories, process querying, and process model matching.

Process model matching is the task of finding correspondences between the activities of two given process models. In particular, for very large process model repositories of organizations, it is essential to utilize process model matching techniques in order to determine similar models and merge them, eliminate redundancies, as well as alleviate storage and processing costs, and increase efficiency accordingly.

Most of the existing model matching techniques typically utilize activity labels and process structures to determine process matching in model repositories [12]. However, incorporating the behavior of the underlying process models is indispensable while detecting process match-ing. Unlike label-based and structure-based process matching approaches, behavioral process model matching takes the order of the activities in the models into consideration to attain a more reliable, accurate matching.

In this paper, we introduce a novel business process model matching approach **Opti**mization-based Process **M**odel M**a**tching (*OPTIMA*) which matches the individual components of two

given process models to each other by enabling the incorporation of both the label and behavioral information of the process models. Our proposal exhibits an optimization problem which maximizes the activity label similarities at an individual local level, and simultaneously maximizes the behavioral similarity of the given processes at a global level by utilizing their *relational profiles* [19, 21, 22, 24]. Thanks to the high flexibility of our approach, it is possible for the user to set the importance (i.e. weighting) of the behavioral information to be incorporated, as well as the label information of the process model components. Furthermore, our approach is completely independent of the application of a prior matching of activity labels, exposing a competitive advantage, when compared with some existing approaches. Moreover, our extensive comparative experimental evaluation performed on three real-world datasets points out the competitiveness of our proposal against the existing techniques, in particular outperforming the state of the art in terms of f-score performance.

Our paper is structured as follows: Section 2 gives an overview of the related work regarding business process model matching. Then, Section 3 presents the preliminaries including fundamental information about Petri nets, as well as relational profiles, and similarity functions we define. In Section 4, we introduce our proposal Optimization-based Process Model Matching (*OPTIMA*), followed by Section 5 which presents the extensive experimental results. Our paper is concluded by Section 6 with a conclusion and future work.

2   Related Work

Business process model matching has been a challenging research area where there have been numerous attempts to provide effective and accurate techniques. Process model matching describes the task of finding corresponding transitions in two given process models, whose roots stem from process model similarity [4, 6, 7, 16, 17] and ontology matching [8] relying on structural and label comparison of processes [2, 5]. Researchers have primarily developed label-based matching techniques which assesses the similarity of acitivity labels in process models. Exhibiting a well-known label-based approach, the basic Bag of Words (BoW) matching technique [11] first determines pairwise bag of words similarity among the labels of transitions, and a word similarity function is used, such as Levenshtein [23] or Lin [13], to compute all pairwise similarity scores and find out the highest scores for the matching.

In contrast to ontology and label-based matching, process models exhibit additional behavioral information which cannot be captured by only considering labels or process structures. Based on this fact, researchers have developed further approaches considering the behavioral information of process models. The authors of [12] propose a behavioral model matching approach which considers both label-based similarities and behavioral relations. After determining the semantic similarity of label components, match constraints are derived based on behavioral profiles [22] of the process models. These constraints are utilized towards a matching formalized as an optimization problem and solved by using Markov Logic Network inference. Another further model matching approach is proposed in [3] which is based on the quantitative bisimulation. First, process models are converted into labeled transition systems and then the degree of simulation is computed, followed by solving a linear program, corresponding to the overall bisimulation result. We refer to [16, 17] for a more comprehensive study of model matching approaches.

Since our proposal can incorporate both the information of activity labels and behavior of given two processes regulated by a parameter, it is noteworthy to use the Bag of Words approach, a label-based method, as a baseline for the label-based matching comparison for our evaluations later on.

3 Preliminaries

For investigating process model matching, we first use Petri nets and workflow nets as our formal grounding [1]. Then, we formulate the relational profile exhibiting a compact behavioral representation of a Petri net. Last, we present similarity functions and give the definition of the relation type similarity we require for our proposal later on.

3.1 Petri Nets

Originally introduced by C. Adam Petri [15], Petri nets are the most utilized process modeling language which allows for concurrency modelling, as well as the analysis of process models effectively. Below, we first present the definition of Petri net, labeled Petri net, and workflow net definitions and terms based on [1], serving as fundamentals for our paper.

Definition 1 (Petri net) *A* Petri net *is defined as a triplet* $N = (P, T, F)$ *where* P *is a finite set of* places, T *a finite set of* transitions *such that* $P \cap T = \emptyset$, *and* $F \subseteq (P \times T) \cup (T \times P)$ *is the* flow relation *denoting a set of directed arcs. A* marked Petri net *is defined as a pair* (N, M) *where* N *is a Petri net and* $M \in \mathbb{B}(P)$ *is a multi-set over* P *denoting the marking of the net.*

Definition 2 (Labeled Petri net) *Let* $\mathcal{A}$ *denote the* universe of activity labels. *A* Labeled Petri net *is a tuple* $N = (P, T, F, A, \lambda)$ *where* (P, T, F) *is a Petri net,* $A \subseteq \mathcal{A}$ *is the set of* activity labels, *and* $\lambda : T \to A$ *is the* labeling function.

For some particular transitions which are not observable, we use the notation τ, i.e. a transition t with $l(t) = \tau$ is unobservable and is referred to as *silent* or *invisible*. Furthermore, elements of $P \cup T$ are referred to as *nodes*. For any $x \in P \cup T$, the *pre-set of* x (a.k.a. input set), denoted $\bullet x$, is the set of nodes with a directed arc to x, i.e. $\bullet x = \{y \mid (y, x) \in F\}$. The *post-set of* x, denoted $x\bullet$, is the set of nodes with a directed arc from x, i.e. $x\bullet = \{y \mid (x, y) \in F\}$.

A marked, labeled Petri net is referred to as *labeled Petri net system*, denoted $S = (N, M_0)$, where $N = (P, T, F, A, \lambda)$ is a labeled Petri net and $M_0 \in \mathbb{B}(P)$ a multi-set over the places P, denoting the *initial marking*. We let $\mathcal{N}$ denote the *universe of marked labeled Petri nets*.

As convention, for any labeled Petri net system $S = (N, M)$ with $N = (P, T, F)$, we let $\mathcal{T}$ denote the *universe of transitions*, and $T_v(S) := \{t \in T \mid \lambda(t) \neq \tau\}$ be the set of non-silent (a.k.a. visible) transitions in S. For sake of simplicity, the notation $T_v(S)$ is replaced by T_S^v in the remainder of the paper, where necessary.

Given a labeled Petri net system (N, M) with $N = (P, T, F, A, \lambda)$, the transition $t \in T$ is *enabled* in marking M, denoted $(N, M)[t\rangle$, iff $\bullet t \leq M$. The *firing rule* $_-[_-\rangle_- \subseteq \mathcal{N} \times T \times \mathcal{N}$ is the smallest relation satisfying for any $(N, M) \in \mathcal{N}$ and any $t \in T : (N, M)[t\rangle \implies (N, M)[t\rangle(N, M) \setminus \bullet t) \uplus t\bullet$.

For a given labeled Petri net system (N, M_0), a sequence $\sigma = \langle t_1, ..., t_n \rangle \in T^*$, with $n \in \mathbb{N}$, is called *firing sequence* of (N, M_0) iff there exist markings $M_1, ..., M_n$ such that for all i with $0 \leq i < n$, t_{i+1} is enabled in marking M_i, i.e. $(N, M_i)[t_{i+1}\rangle$, and firing t_{i+1} ends up in the marking M_{i+1}, i.e. $(N, M_i)[t_{i+1}\rangle(N, M_{i+1})$.

Workflow nets, a subclass of Petri nets, are highly relevant for business process modeling due to their strength in natural representation of the life-cycle of cases of the underlying process models [1]. The formal definition of workflow net is given below.

Definition 3 (Workflow net) *Given an identifier* $\bar{t} \notin P \cup T$, *a labeled Petri net* $N = (P, T, F, A, \lambda)$ *is called a* workflow net *(WF-net) iff (1)* P *contains a source place* $i \in P$ *(a.k.a. input place) such that* $\bullet i = \emptyset$, *(2)* P *contains a sink place* $o \in P$ *(a.k.a. output place) such that* $o\bullet = \emptyset$ *and (3) its short circuit net* $\bar{N} = (P, T \cup \{\bar{t}\}, F \cup \{(o, \bar{t}), (\bar{t}, i)\}, A \cup \{\tau\}, \lambda \cup \{(\bar{t}, \tau)\})$ *is strongly connected, i.e. there is a directed path between any pair of nodes in* $\bar{N}$.
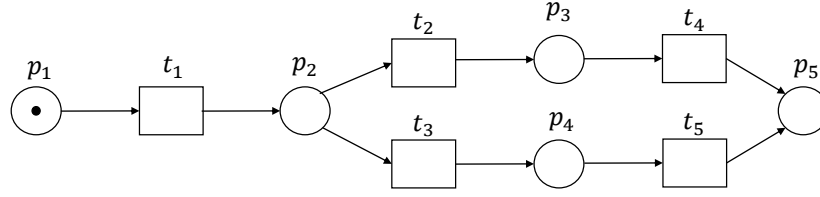
Figure 1. An example workflow net. The notation p_i denotes the i-th place and t_j denotes the j-th transition. The places p_1 and p_5 exhibit the input (aka source) place and output (aka sink) place, respectively.

Since WF-nets can expose processes with errors, such as deadlocks, activities that can never become active, still enabled intermediate transitions in spite of the process termination, etc., we need to define soundness criterion which is commonly used in the literature [20].

A workflow net $N = (P, T, F, A, \lambda)$ with an input place $i \in P$ and an output place $t \in P$ is called *sound* iff (1) $(N, [i])$ is safe, i.e. places cannot hold multiple tokens at the same time (*safeness*), (2) for any marking $M \in [N, [i]\rangle$: $o \in M \Rightarrow M = [o]$ (*proper completion*), (3) for any marking $M \in [N, [i]\rangle : [o] \in [N, M\rangle$ (*option to complete*), (4) for any transition $t \in T$, there is a firing sequence enabling t, i.e. $(N, [i])$ includes no dead transitions (*absence of dead parts*). Furthermore, a Petri net is *free-choice* if any two transitions sharing an input place have identical input sets, i.e. for all transitions $t_1, t_2 \in T, \bullet t_1 \cap \bullet t_2 \neq \emptyset \Rightarrow \bullet t_1 = \bullet t_2$. Figure 1 exhibits an example workflow net.

3.2 Relational Profiles

In order to give a compact behavioral representation of a Petri net, an appropriate structure is required which captures the relationships among its transitions. Below, we present the comprehensive definition of the *relational profile*.

Definition 4 (Relational profile) *Let $N = (P, T, F, A, \lambda)$ be a sound free-choice workflow net and $S = (N, M_0)$ the corresponding workflow net system. A* relational profile $\mathcal{R}^S = (\Psi, \Omega)$ *of S is a tuple comprising a set Ψ of* relation types *and an* assignment relation $\Omega \subseteq T \times T \times \Psi$ *which assigns pairs of transitions relation types. A transition $s \in T$ is in a relation $R \in \Psi$ with a transition $t \in T$, denoted sRt, iff $(s, t, R) \in \Omega$. $\mathcal{R}^S$ is called* mutually exclusive relational profile *if for all transitions $s, t \in T$ and all relation types $R_1, R_2 \in \Psi$ with $R_1 \neq R_2 : (s, t, R_1) \in \Omega \Rightarrow (s, t, R_2) \notin \Omega$.*

Since our proposal *OPTIMA* requires that profiles assign at most one relation per pair of transitions, we will consider such relational profiles satisfying the latter via the term *mutually exclusive profiles*, complying with [21].

Example. We consider a relational profile $\mathcal{R}^S = (\Psi, \Omega)$ of the workflow net in Figure 1 and two relation types *eventually-follows relation* $\succ \subseteq T \times T$ and *directly-follows relation* $> \subseteq T \times T$, resulting in $\Psi = \{\succ, >\}$. Note that (t_i, t_j) is in an eventually-follows relation if there exists a firing sequence which fires t_i before t_j. In contrast, (t_i, t_j) is in a directly-follows relation if there exists a firing sequence where t_j is fired after t_i without any visible transition in between. In Figure 1, we realize that $t_1 \succ t_4$ holds but t_1 is not directly-followed by t_4, i.e. $t_1 \not> t_4$, thus, $(t_1, t_4, \succ) \in \Omega$ and $(t_1, t_4, >) \notin \Omega$. In addition, $\mathcal{R}^S$ is not a mutually exclusive relational profile, since two transitions can exhibit more than one relation, e.g. $(t_1, t_2, \succ) \in \Omega$ and $(t_1, t_2, >) \in \Omega$.

3.3 Similarity Functions

After presenting the definition of *relational profile*, we now focus on the similarity computation of two relational profiles of the Petri nets at hand. Since we assume that mutually exclusive profiles are used to represent the behavior of Petri nets, we define a similarity function which

determines the similarity of two given relation types, corresponding to the behavioral similarity of two transitions exposing those relation types, such as directly-follows relation and eventually-follows relation. In this paper, since we consider the relational profiles of the α-relational Profile ($\alpha\mathcal{P}$) [19], the Behavioral Profile ($\mathcal{BP}$) [21, 22], and the BP+ profile ($\mathcal{BPP}$) [24], we define the *relation type similarity* function by using the aforementioned profiles. Please note that this similarity function is not limited to these profiles, and can easily be extended by the further profiles accordingly, where necessary.

Definition 5 (Relation type similarity) *Let S_1 and S_2 be sound and free-choice WF-net systems with relational profiles $\mathcal{R}^{S_1} = (\Psi, \Omega_1)$ and $\mathcal{R}^{S_2} = (\Psi, \Omega_2)$ of type $\mathcal{R} \in \{\mathcal{BP}, \alpha\mathcal{P}, \mathcal{BPP}\}$. The relation type similarity $sim^{\mathcal{R}} : \Psi \times \Psi \to [0, 1]$ of two relation types $R_1, R_2 \in \Psi$ is defined by:*

$$sim^{\mathcal{R}}(R_1, R_2) = \begin{cases} \mathbb{1}[R_1 = R_2] & \text{if } \mathcal{R} \in \{\mathcal{BP}, \alpha\mathcal{P}\} \\ w_{R_1, R_2} & \text{if } \mathcal{R} = \mathcal{BPP} \end{cases}$$

where the similarity value w_{R_1, R_2} of BP+ relation types stems from [24](Table 2).

The identification function $\mathbb{1}[\alpha] \in \{0, 1\}$ returns 1 if and only if the statement α is true, i.e. if the relation equivalence holds in the definition above.

Analogously, the *label-based similarity* function $sim^L : \mathcal{T} \times \mathcal{T} \to [0, 1]$ computes the similarity of the given two transitions, which will be utilized in the upcoming section, too.

4 Optimization-based Process Model Matching

In this section, we propose our novel approach **Opti**mization-based Process **M**odel M**a**tching (*OPTIMA*) which takes both local and global information of the underlying process models into consideration. This is achieved by utilizing the label information of the activity labels and the behavior information of both process models.

Our approach is presented as an optimization problem which maximizes the label similarities at an individual local level, and simultaneously maximizes the behavioral similarity of both processes at a global level by using their relational profiles. This is attained by defining an integer linear program which exhibits an optimization problem with a linear objective function, linear constraints, and variables which are defined to be integers [10].

In order to provide flexibility for the user, e.g. process owner, domain expert, etc., we introduce a weighting parameter w which determines how much importance will be attached to label information and behavioral information, aligning with the user intention. Moreover, our proposal is fully independent of the application of a prior matching of transition labels, which constitutes an important competitive advantage in comparison with some existing approaches. For sake of simplicity, the notations $T_v(S_1)$ and $T_v(S_2)$ will be replaced by T_1^v and T_2^v for the remainder of our paper, where required. Below, we first give the formal definition of our novel approach *OPTIMA* and then elaborate on its constraints:

Definition 6 (Optimization-based Model Matching) *Given two sound free-choice WF-net systems $S_1 = (N_1, M_0^{S_1})$, $S_2 = (N_2, M_0^{S_2})$ with $N_1 = (P_1, T_1, F_1, A_1, \lambda_1)$, $N_2 = (P_2, T_2, F_2, A_2, \lambda_2)$, and mutually exclusive relational profiles $\mathcal{R}^{S_1} = (\Psi, \Omega_1)$ of S_1 and $\mathcal{R}^{S_2} = (\Psi, \Omega_2)$ of S_2, let $sim^{\mathcal{R}} : \Psi \times \Psi \to [0, 1]$ be a relation type similarity of the profile type $\mathcal{R}$ and $sim^L : \mathcal{T} \times \mathcal{T} \to [0, 1]$ be a label-based similarity function. The Optimization-based Model Matching (OPTIMA) $M \subseteq T_1^v \times T_2^v$ is derived from the optimal solution of the following problem:*

$$\max \ w \sum_{\substack{s_1,s_2 \in T_1^v \\ t_1,t_2 \in T_2^v}} \frac{1}{m^2} y_{s_1,s_2,t_1,t_2} sim^{\mathcal{R}}(\mathcal{R}_{s_1,s_2}^{S_1}, \mathcal{R}_{t_1,t_2}^{S_2}) \ + \ (1-w) \sum_{\substack{s \in T_1^v \\ t \in T_2^v}} \frac{1}{m} x_{s,t} sim^L(s,t)$$

$$\text{s.t.} \quad \sum_{s \in T_1^v} x_{s,t} \leq 1 \qquad\qquad\qquad \forall t \in T_2^v \qquad (1)$$

$$\sum_{t \in T_2^v} x_{s,t} \leq 1 \qquad\qquad\qquad \forall s \in T_1^v \qquad (2)$$

$$2y_{s_1,s_2,t_1,t_2} \leq x_{s_1,t_1} + x_{s_2,t_2} \qquad \forall s_1, s_2 \in T_1^v, t_1, t_2 \in T_2^v \qquad (3)$$

$$x_{s,t} \in \{0,1\} \qquad\qquad\qquad \forall s \in T_1^v, t \in T_2^v \qquad (4)$$

$$y_{s_1,s_2,t_1,t_2} \in \{0,1\} \qquad\qquad \forall s_1, s_2 \in T_1^v, t_1, t_2 \in T_2^v \qquad (5)$$

where $w \in [0,1]$ denotes the weighting parameter, and $m = \min\{|T_1^v|, |T_2^v|\}$.

The maximum number of simple correspondences of the two nets N_1 and N_2, i.e. the matching of single transitions of Petri nets, is determined by $m := \min\{|T_1^v|, |T_2^v|\}$. According to Constraint (4) above, for transitions $s \in T_1^v$ and $t \in T_2^v$, $x_{s,t} \in \{0,1\}$ indicates if s is matched to t (i.e. $x_{s,t} = 1$) or not (i.e. $x_{s,t} = 0$). Constraints (1) and (2) ensure that every transition of one WF-net is matched to at most one transition of the other WF-net.

The decision variable y is concerned with the aggregation of the information of two x variables: For visible transitions $s_1, s_2 \in T_1^v$ and $t_1, t_2 \in T_2^v$, Constraint (5) indicates if s_1 is matched to t_1 and simultaneously if s_2 is matched to t_2 (i.e. $y_{s_1,s_2,t_1,t_2} = 1$) or not (i.e. $y_{s_1,s_2,t_1,t_2} = 0$). Furthermore, Constraint (3) denotes the relationship between the variables x_{s_1,t_1}, x_{s_2,t_2}, and y_{s_1,s_2,t_1,t_2} which ensures that if $x_{s_1,t_1} = x_{s_2,t_2} = 1$, then the maximization problem results in $y_{s_1,s_2,t_1,t_2} = 1$ due to the nature of the problem definition.

The objective function comprising two summands aims to maximize the average label similarity between matched transitions, and maximize the behavioral similarity of both WF-nets, depending on their relational profiles. Finally, the obtained sum is normalized by the squared number of maximum possible simple correspondences m^2 to provide an objective value in $[0.0, 1.0]$.

5 Experiments

In this section, we first give details about the experimental system setup, datasets, and the process model matching approaches which are used in our evaluations. Then, we will present the extensive evaluation results.

5.1 Experimental Setup

5.1.1 System setup.

The implementation of programs is performed in JAVA 8 and experiments are conducted on 2 × Intel Xeon Gold 5115 CPUs, each consisting of 10 cores and 20 threads @ 2.40GHz with a total of 512 GB RAM DDR4-2400 and 15 × 400-AXQU 960 GB SSD with Ubuntu Linux 18.04. In addition, for our proposed *OPTIMA* approach, we utilize Gurobi 8.0.1 [9], and adopt the Petri net and behavioral profile implementation from the jBPT[1] library. The implementation utilized for the evaluation results presented in this paper is available and can be publicly checked out[2].

5.1.2 Datasets.

We use three real-world datasets which arise from the Process Model Matching Contests 2015 [2]. The first dataset is the University Admission Processes (abbreviated by *University*) comprising 36 model pairs derived from 9 models representing the application procedure for students

---

[1]https://github.com/jbpt/codebase
[2]https://github.com/domhues/ilp-matcher

	Characteristics	*University*	*Birth*	*Asset*
Before conversion to PNML	# model pairs	36*	36	36*
	# transitions (min)	12*	9	1*
	# transitions (max)	45*	25	43*
	# transitions (avg)	24.2*	19.3	18.6*
After conversion to PNML	# model pairs	21	-	17
	# non-silent transitions (min)	16	-	1
	# non-silent transitions (max)	32	-	21
	# non-silent transitions (avg)	24.1	-	6.4
Gold standard	% matched non-silent transitions	25.35%	65.95%	84.86%
	% unmatched non-silent transitions	74.65%	34.05%	15.14%
	% simple correspondences	83.3%	14.0%	22.0%
	% complex correspondences	16.7%	86.0%	78.0%
	% trivial correspondences	33.3%	4.0%	18.3%

Table 1. Characteristic information of the datasets *Birth, University,* and *Asset* ($*$ values are adopted from [2]).

at nine universities in Germany. The second dataset is the Birth Registration Processes (*Birth*) consisting of 36 model pairs that were derived from 9 models representing the birth registration processes of Germany, Russia, South Africa, and the Netherlands. The third dataset is Asset Management Processes (*Asset*) which includes 36 model pairs that were derived from 72 models from an SAP Reference Model Collection covering the fields of finance and accounting. Since the *University* and *Asset* datasets originally include process models of BPMN and EPML formats, respectively, these models are first converted to Petri nets, i.e. PNML format, so that process model matching approaches and our proposal can be evaluated.

It is noteworthy to state that model pairs available in the datasets *Birth, University,* and *Asset* are associated with a *gold standard* indicating the ground truth corresponding to the optimal matching of the process model pairs. The gold standard is derived manually by making use of the human expert knowledge, comprising simple (*1:1*) and complex (*1:n*) correspondences.

Table 1 presents key characteristics about the three aforementioned datasets. As mentioned above, the process models in the datasets *University* and *Asset* are of BPMN and EPML formats, respectively, which are converted to Petri nets, i.e. PNML format. It is visible that the conversion of the models from BPMN and EPML into PNML format affects the number of model pairs which are then used for the matching evaluations. The reason for obtaining less process models after the format conversion is that some transformed models are not free-choice models any more to which relational profiles cannot be applied. Since the *Birth* dataset comprises the process models of the PNML format, there is no need to apply any model conversion, i.e. 36 process model pairs remain for the experimental evaluation. In addition, the *University* dataset indicates the highest number of non-silent transitions (24.1 transitions), while the *Asset* dataset has the lowest number of non-silent transitions (6.4 transitions). Furthermore, the *Asset* dataset includes minimum 1 non-silent transition, while the *University* dataset shows the highest minimum number of non-silent transitions (16).

At the bottom of Table 1, we realize some information regarding the gold standard indicating the percentage of the matched and unmatched transitions, as well as the information of simple and complex correspondences. For the *University* dataset, 25.35% of the non-silent transitions are considered as mapped by the gold standard, while 74.65% remain unmatched. Furthermore, the *University* dataset indicates a high percentage of simple correspondences (83.3%) in the gold standard, while the *Asset* dataset shows a high percentage of complex correspondences (78%) in its gold standard. Please note that a low percentage of simple correspondences corresponds to a high percentage of complex correspondences.

5.1.3 Approaches.

As given above, our proposal *OPTIMA* utilizes both label-based and behavioral information of process models by means of a weighting parameter w. We vary $w \in \{0.0, 0.1, ..., 1.0\}$ and utilize the basic Bag of Words label similarity function with Lin word similarity function. Furthermore, we consider the following individual relational profiles of process models: the α-relational Profile ($\alpha\mathcal{P}$) [19], the Behavioral Profile ($\mathcal{BP}$) [21, 22], and the BP+ profile ($\mathcal{BPP}$) [24]. Please note that the evaluation and comparison of various label similarity functions is not the scope of this work.

In order to provide a fair empirical study, we consider the existing process model matching approaches which in particular take the behavioral information of process models into consideration, too. First, we utilize the Markov Logic Network model matching approach [12] with two different labeling functions, namely **R**efactored label similarity function, as proposed by the authors, and the basic **B**ag of Words label similarity function with Lin word similarity function [13] (abbreviated by *MarkovR* and *MarkovB*, respectively) so that the results can be compared with those of *OPTIMA*. Then, we use the bisimulation-based model matching approach (*Bisim*) [3] utilizing the basic Bag of Words label similarity function with Lin word similarity function so that the results are comparable with those of our proposal.

It is noteworthy to state that setting $w = 0$ leads to the fact that *OPTIMA* exposes only a label-based process model matching, while setting $w = 1$ provides our proposal to include only behavioral information for process model matching. Hence, regarding the former, we consider a baseline from the class of the label-based process model matching approaches, i.e. the Bag of Words process model matching approach [11] (*BoW*) with Lin word similarity function (with the threshold value of 0.7, as authors suggest).

We compare our computed matching results, i.e. found correspondences in our matching, against a gold standard which is generated by authors in [2]. In this way, each found correspondence of the activity pair is determined to be in one of the following classes: true-positive (TP), true-negative (TN), false-positive (FP) or false-negative (FN). Taking this classification into consideration, we then calculate the precision ($TP/(TP + FP)$), the recall ($TP/(TP + FN)$), and the f-score $(2 \times precision \times recall)/(precision + recall)$.

In order to gather comparable, fair evaluation results, we utilize the average of precision, recall, and f-score values, referring to [2]: the *macro* evaluation considers the average of precision, recall, and f-score values over all test cases, while the *micro* evaluation is obtained by first summing up all true/false positives, true/false negatives, and then computing the precision, recall, and f-score values once at the end of the computation. Furthermore, since *Bisim*, *MarkovR, MarkovB*, and our proposal *OPTIMA* consist of parameters which are not necessarily preset by the authors, we present our results obtained by applying the parameters which lead to the highest micro f-score values, as stated in [2]. Due to space limitations, we focus on the precision, recall, and f-score comparison of our proposal and the aforementioned approaches.

5.2 Experimental Results

After giving details about the experimental setup above, we now present our evaluation results. Based on *micro* and *macro* aggregation of the results, Figure 2 exhibits precision, recall, and f-score results of our proposal *OPTIMA* and four state-of-the-art approaches, respectively. By inspecting the plots in both figures, we note that micro and macro aggregated evaluation results of all approaches expose a similar tendency over utilized real-world datasets (*Birth, Asset, and University*), as anticipated.

As stated before, we present the results attained by applying the parameters resulting in the highest micro f-score values, as proposed in [2]. The results of our approach *OPTIMA* are obtained by determining the best values attained with the BP+ profile ($\mathcal{BPP}$) with the weighting parameter $w = 0.4$ on the *Birth* dataset, with the BP+ profile ($\mathcal{BPP}$) with the weighting param-

(a) Precision results obtained by *macro* aggregation.

(b) Precision results obtained by *micro* aggregation.

(c) Recall results obtained by *macro* aggregation.

(d) Recall results obtained by *micro* aggregation.

(e) F-score results obtained by *macro* aggregation.
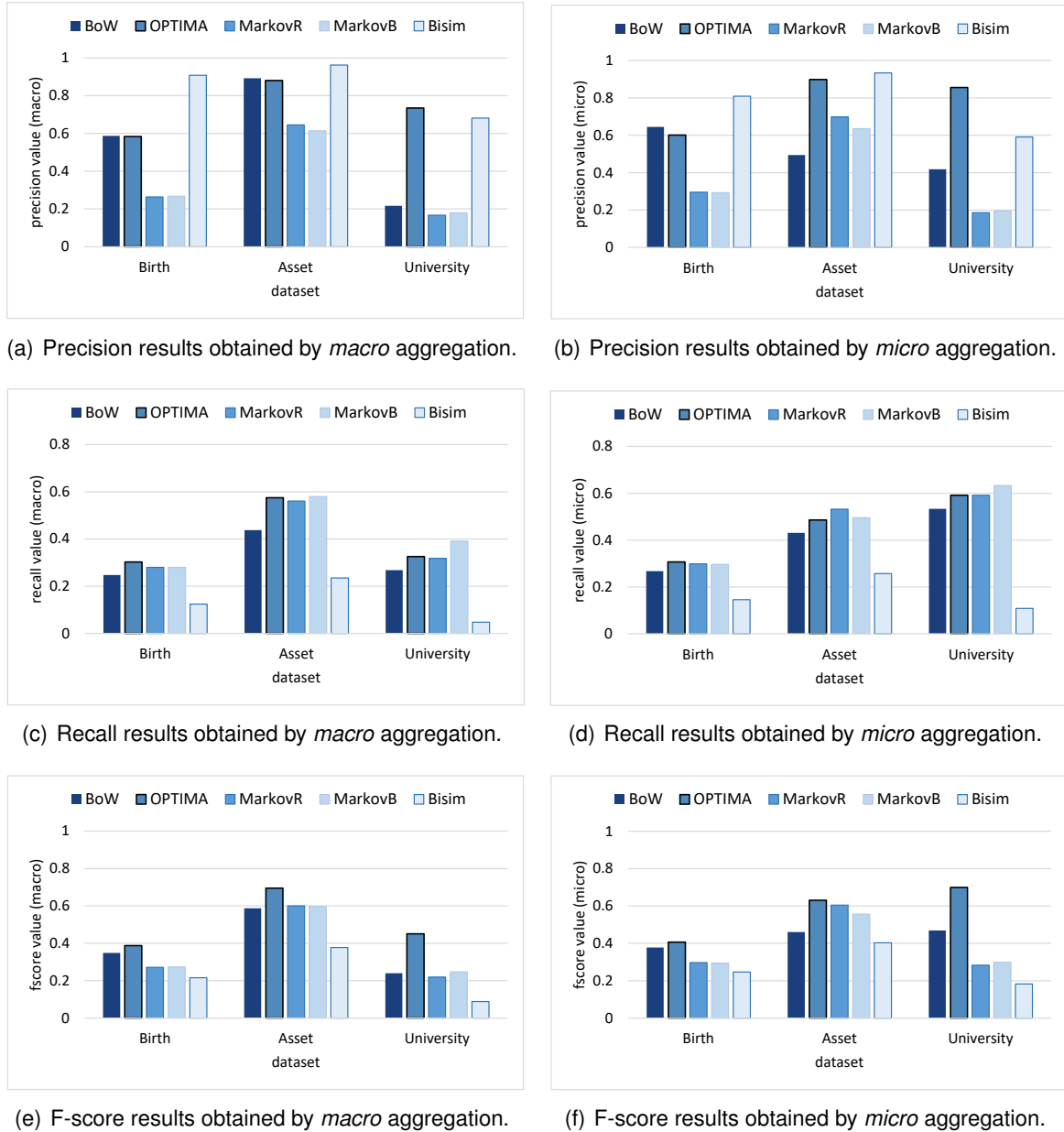
(f) F-score results obtained by *micro* aggregation.

Figure 2. *Micro* and *macro* aggregation results of precision, recall, and f-score measures obtained on three real-world datasets, i.e. Birth Registration Processes (*Birth*), Asset Management Processes (*Asset*), and University Admission Processes (*University*) stemming from Process Model Matching Contest 2015 [2]. For a fair comparison, we utilize (i)Bag of Words process model matching approach [11] (*BoW*) indicating a baseline for a label-based approach (ii) Markov Logic Network model matching approach [12] with two variations *MarkovR* and *MarkovB* utilizing Refactored and Bag of Words label similarity functions (iii) Bisimulation-based model matching approach [3] (*Bisim*). Considering both label and behavior information of process models and being independent of the application of a prior matching of activity labels, our proposal *OPTIMA* outperforms all approaches regarding micro and macro f-score results on all real-world datasets.

eter $w = 0.5$ on the *University* dataset, and with the Behavioral Profile ($\mathcal{BP}$) with the weighting parameter $w = 0.2$ on the *Asset* dataset. Furthermore, for the bisimulation-based approach *Bisim*, a skip-penalty with the value 0.7 for *Birth*, *Asset*, and a skip-penalty with the value 0.9 for *University* lead to the highest values. In addition, the best values for Markov-based approaches *MarkovB*, *MarkovR* are attained by the constraint weight 0.001 for *Birth*, *University*, and the con-

straint weight 0.01 for the *Asset* dataset. Finally, for the Bag of Words process model matching approach (*BoW*) indicating a baseline for a label-based matching model technique, we utilize the threshold of 0.7 which is suggested by the authors in [11].

The results summarized in Figure 2(a)-2(b) report that the *Bisim* approach outperforms other approaches in terms of precision performance on the *Birth* and *Asset* datasets. Furthermore, *OPTIMA* exhibits the highest precision values on the *University* dataset, while the Markov-based approaches show the lowest precision performance. The slightly higher performance of *Bisim* over that of *OPTIMA* can be elucidated by the fact that the quantitative simulation exposes a higher or comparable expressiveness for model matching, when compared with the incorporation of the relational profiles. Furthermore, we observe that the *BoW* approach, indicating only a baseline label-based approach, shows a much higher precision performance than that of *MarkovB* and *MarkovR* on *Birth* and *Asset* comprising complex correspondences in their gold standard, i.e. ground truth. This can be elucidated by the fact that *BoW* can successfully detect complex correspondences, while Markov-based approaches can only find a smaller portion of complex correspondences on these datasets.

The results presented in Figure 2(c)-2(d) provide confirmatory evidence that our proposal *OPTIMA* outperforms existing approaches regarding the recall measure on the *Birth* dataset. In addition, *OPTIMA* indicates a comparable recall performance when compared with *MarkovR* and *MarkovB* on all datasets, which can be explained by that fact that both Markov-based approaches and our proposal can detect simple correspondences successfully. Furthermore, *BoW* shows a comparable recall performance on all datasets, outperforming *Bisim*. An interesting observation is the considerably poor recall performance of *Bisim* on all datasets. This suggests that the applied quantitative simulation technique is eligible for detecting only a small fraction of the relevant results.

A closer examination of macro and micro aggregation indicates that the macro aggregation of the precision, recall, and f-score measures reflects lower values on the *University* dataset than the micro aggregation results. This posits that some model pairs in *University* expose a relatively poor performance in the three measures, directly resulting in lower macro aggregated values since every model matching contributes equally to the computation of macro scores. In contrast, the corresponding micro aggregation results seem to be higher, since the influence of the poor performance of some particular models in the aforementioned dataset does not substantially contribute to the computation of micro aggregation results at all.

As presented in Figure 2(e)-2(f), *OPTIMA* considerably outperforms all state-of-the-art matching approaches regarding both micro and macro aggregation f-score results on all three real-world datasets. The intuition behind this observation lies in the high precision and recall results of our proposed approach, while other approaches exhibit a smaller result either in precision or recall.

6 Conclusion

One of the major challenges and key components in today's organizations is the ever-increasing amounts of business processes, resulting in the need for novel effective process model matching techniques in huge process model repositories. Providing direct insight into process model matching, this paper introduces a novel business process model matching approach *Optimization-based Process Model Matching* (*OPTIMA*) which matches individual components of two given process models to each other by incorporating both label and behavioral information of the process models. We present an optimization problem maximizing the activity label similarities at a local level, and the behavioral similarity of the given processes at a global level by leveraging relational profiles. Being fully independent of any prior matching of activity labels, our proposal shows high competitiveness against existing techniques, in particular outperforming the state

of the art in terms of accuracy performance.

An interesting direction for future work concerns the analysis of the complex correspondences which can potentially shed on light on the further matching strategies. Furthermore, we intend to conduct research into the evaluations on various real-world datasets in order to gain more insights, as well as examine the execution time of our approach to evaluate its performance. In addition, examining the efficiency of the proposal, as well as reducing the number of variables in the maximization problem to attain higher efficiency can be dedicated for future examination.

Acknowledgments

References

[1] van der Aalst, W.M.P.: Process Mining: Data Science in Action. Springer, Heidelberg, 2 edn. (2016)

[2] Antunes, G., Bakhshandeh, M., Borbinha, J., Cardoso, J., Dadashnia, S., Francescomarino, C.D., Dragoni, M., Fettke, P., Gal, A., Ghidini, C., Hake, P., Khiat, A., Klinkmüller, C., Kuss, E., Leopold, H., Loos, P., Meilicke, C., Niesen, T., Pesquita, C., Peus, T., Schoknecht, A., Sheetrit, E., Sonntag, A., Stuckenschmidt, H., Thaler, T., Weber, I., Weidlich, M.: The Process Model Matching Contest 2015. In: EMISA'15: International Workshop on Enterprise Modelling and Information Systems Architecture. pp. 127–155. GI, Innsbruck, Austria (Sep 2015)

[3] Becker, J., Breuker, D., Delfmann, P., Dietrich, H.A., Steinhorst, M.: Identifying Business Process Activity Mappings by Optimizing Behavioral Similarity. In: AMCIS. vol. 1, p. Paper 21 (01 2012)

[4] Becker, M., Laue, R.: A Comparative Survey of Business Process Similarity Measures. Computers in Industry **63**(2), 148 – 167 (2012)

[5] Cayoglu, U., Dijkman, R., Dumas, M., Fettke, P., García-Bañuelos, L., Hake, P., Klinkmüller, C., Leopold, H., Ludwig, A., Loos, P., Mendling, J., Oberweis, A., Schoknecht, A., Sheetrit, E., Thaler, T., Ullrich, M., Weber, I., Weidlich, M.: Report: The Process Model Matching Contest 2013. In: Lohmann, N., Song, M., Wohed, P. (eds.) Business Process Management Workshops. pp. 442–463. Springer International Publishing, Cham (2014)

[6] Dijkman, R.M., Dumas, M., van Dongen, B.F., Käärik, R., Mendling, J.: Similarity of Business Process Models: Metrics and Evaluation. Inf. Syst. **36**(2), 498–516 (2011)

[7] Dumas, M., García-Bañuelos, L., Dijkman, R.M.: Similarity Search of Business Process Models. IEEE Data Eng. Bull. **32**(3), 23–28 (2009)

[8] Euzenat, J., Shvaiko, P.: Ontology Matching. Springer Publishing Company, Incorporated, 2nd edn. (2013)

[9] Gurobi Optimization LLC: Gurobi Optimizer Reference Manual (2019), `http://www.gurobi.com`

[10] Hillier, F., Lieberman, G.: Introduction to Linear Programming. McGraw-Hill (1990)

[11] Klinkmüller, C., Weber, I., Mendling, J., Leopold, H., Ludwig, A.: Increasing Recall of Process Model Matching by Improved Activity Label Matching. In: Business Process Management. pp. 211–218. Springer Berlin Heidelberg (2013)

[12] Leopold, H., Niepert, M., Weidlich, M., Mendling, J., Dijkman, R., Stuckenschmidt, H.: Probabilistic Optimization of Semantic Process Model Matching. In: Business Process Management. pp. 319–334. Springer Berlin Heidelberg (2012)

[13] Lin, D.: An Information-theoretic Definition of Similarity. In: Proc. of the 15th International Conference on Machine Learning. vol. 98, pp. 296–304. Morgan Kaufmann (1998)

[14] Pegoraro, M., Uysal, M.S., van der Aalst, W.M.P.: Discovering Process Models from Uncertain Event Data. In: Business Process Management Workshops. pp. 238–249. Springer International Publishing, Cham (2019)

[15] Petri, C.A.: Kommunikation mit Automaten. Schriften des Rheinisch-Westfälischen Institutes für Instrumentelle Mathematik an der Universität Bonn, Technische Hochschule, Darmstadt. (1962), `https://books.google.de/books?id=NCZMvAEACAAJ`

[16] Schoknecht, A., Thaler, T., Fettke, P., Oberweis, A., Laue, R.: Similarity of Business Process Models – A State-of-the-Art Analysis. ACM Comput. Surv. **50**(4), 52:1–52:33 (Aug 2017)

[17] Thaler, T., Schoknecht, A., Fettke, P., Oberweis, A., Laue, R.: A Comparative Analysis of Business Process Model Similarity Measures. In: Business Process Management Workshops. pp. 310–322. Springer International Publishing, Cham (2017)

[18] Uysal, M.S., van Zelst, S.J., Brockhoff, T., Ghahfarokhi, A.F., Pourbafrani, M., Schumacher, R., Junglas, S., Schuh, G., van der Aalst, W.M.: Process Mining for Production Processes in the Automotive Industry. In: Industry Forum at BPM 2020 co-located with 18th International Conference on Business Process Management (BPM 2020), Sevilla, Spain (2020)

[19] van der Aalst, W., Weijters, T., Maruster, L.: Workflow Mining: Discovering Process Models from Event Logs. IEEE Transactions on Knowledge and Data Engineering **16**(9), 1128–1142 (Sep 2004)

[20] van der Aalst, W.: The Application of Petri-nets to Workflow Management. Journal of Circuits, Systems and Computers **8**(1), 21–66 (1998)

[21] Weidlich, M., Mendling, J., Weske, M.: Efficient Consistency Measurement Based on Behavioral Profiles of Process Models. IEEE Transactions on Software Engineering **37**(3), 410–429 (May 2011)

[22] Weidlich, M., Mendling, J., Weske, M.: Computation of Behavioural Profiles of Process Models. Business Process Technology, Hasso Plattner Institute for IT-Systems Engineering. Potsdam (2009)

[23] Weigel, A., Fein, F.: Normalizing the Weighted Edit Distance. In: Proceedings of the 12th IAPR International Conference on Pattern Recognition, Vol. 3 - Conference C: Signal Processing. vol. 2, pp. 399–402 vol.2 (Oct 1994)

[24] Wen, L., Song, J., Wang, J., Kumar, A.: BP+: An Improved Behavioral Profile Metric for Process Models. https://www.researchgate.net/publication/286932844 (2015), accessed: 01.02.2021