Cross-Department Collaborative Healthcare Process Model Discovery From Event Logs

Cong Liu, *Member, IEEE*, Huiling Li, Shuaipeng Zhang, Long Cheng, *Senior Member, IEEE*, and Qingtian Zeng

***Abstract*— Healthcare plays an increasingly essential role in our daily life. Modern Hospital Information Systems (HISs) record and store detailed medical treatment process information for all patients as event logs. By taking event logs as input, process mining techniques have been widely applied to extract valuable insights to improve medical treatment processes and deliver better healthcare services. However, considering the complexity of collaborations among different medical departments, existing model discovery techniques cannot be applied directly. To handle this limitation, this paper proposes a novel approach to support the discovery of Cross-department Collaborative Healthcare Process (CCHP) models from medical event logs. Specifically, an extension of classical Petri Nets with message and resource attributes is first introduced to formalize CCHPs. Then, a novel discovery algorithm is proposed to discover Intra-department Healthcare Process (IHP) models. Next, collaboration patterns among medical departments are formalized and corresponding discovery algorithms are given on that basis. Finally, a global CCHP model is obtained by integrating all discovered collaboration patterns and IHP models. By using four public medical event logs, we quantitatively compare our approach with the state-of-the-art process mining techniques in terms of model quality, and our experimental results demonstrate that the proposed approach can discover more accurate healthcare process models.**

***Note to Practitioners*—The recorded medical event logs by HISs can be used to extract valuable insights for the analysis of healthcare processes. However, existing process model discovery techniques cannot be applied for the analysis directly due to the complex collaborations among different medical departments of a hospital. This paper introduces a novel approach for cross-department collaborative healthcare process model discovery from medical event logs. All proposed techniques are fully implemented and publicly available. Using four public medical event logs, we show the applicability and advantages of our approach against existing ones. The proposed techniques are applicable to the model discovery and behavior understanding of real-life operational healthcare processes.**

***Index Terms*—Petri nets, model discovery, cross-department healthcare processes, collaboration patterns.**

I. Introduction

WITH the continuous improvement of people's living standards, healthcare services are receiving more and more attentions in the past decades. A healthcare process is a series of medical activities focusing on diagnosing, treating, and preventing diseases for patients [1]. Typically, multiple collaborative departments of a hospital are involved during the execution of healthcare treatment processes, and therefore, such processes are known as Cross-department Collaborative Healthcare Processes (CCHPs). Different from traditional collaborative business processes where processes interact via messages, interactions among CCHPs are more complex. More specifically, besides message interaction patterns, resource sharing patterns and task synchronization patterns are commonly seen. On the one hand, the use of CCHPs can significantly improve patients' treatment experiences and reduce repetitive tests. On the other hand, it allows hospital managers to uncover potential bottlenecks, simplify the medical process, and improve the efficiency of medical treatment services. However, a precise and up-to-date CCHP description is typically not available in real-life cases because of the out-dated or incomplete documentation. Considering the complexity of operational healthcare processes, building CCHP models manually from scratch is typically time-consuming and requires modellers to be with extensive domain knowledge.

With the development of medical digitalization, medical treatment services have been supported by various Hospital Information Systems (HISs). During the execution of HISs, huge amounts of medical event logs will be accumulated and stored, which actually can provide valuable information

Manuscript received 11 January 2022; revised 5 April 2022; accepted 3 May 2022. Date of publication 3 August 2022; date of current version 3 July 2023. This article was recommended for publication by Associate Editor V. Augusto and Editor X. Xie upon evaluation of the reviewers' comments. This work was supported in part by the National Natural Science Foundation of China under Grant 61902222, in part by the Taishan Scholars Program of Shandong Province under Grant ts20190936 and Grant tsqn201909109, in part by the Natural Science Excellent Youth Foundation of Shandong Province under Grant ZR2021YQ45, and in part by the Youth Innovation Science and Technology Team Foundation of Shandong Higher School under Grant 2021KJ031. *(Corresponding authors: Huiling Li; Qingtian Zeng.)*

Cong Liu is with the School of Computer Science and Technology, Shandong University of Technology, Zibo 255000, China, and also with the College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China (e-mail: liucongchina@sdust.edu.cn).

Huiling Li is with the School of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255000, China (e-mail: 17864309765@163.com).

Shuaipeng Zhang is with the School of Computer Science and Technology, Shandong University of Technology, Zibo 255000, China (e-mail: 15994069715@163.com).

Long Cheng is with the School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China (e-mail: lcheng@ncepu.edu.cn).

Qingtian Zeng is with the College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China (e-mail: qtzeng@163.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TASE.2022.3194312.

Digital Object Identifier 10.1109/TASE.2022.3194312

for healthcare process analysis and enhancement, and *process mining* is such a technique which can be used to extract valuable insights from event logs [2]. As one of the most promising process mining techniques, model discovery is able to construct a descriptive model by taking as input an event log. However, existing model discovery techniques, e.g., *Alpha Miner* [3], *Inductive Miner* [4], and *Split Miner* [5], cannot be applied directly to discover precise CCHP models. There are two main reasons: (1) they focus on control-flow information and are unable to handle message and resource factors which are two vital factors in a collaborative healthcare setting; and (2) they focus on intra-department business process while the discovery of cross-department collaboration behavior, e.g., message exchange, resource sharing, and task synchronization, are not properly supported.

To handle the limitations, a novel automated discovery framework is introduced in this work. Specifically, classical Petri nets that are extended with resource and message attributes, denoted as *RM_WF_nets* [6], are used to formalize CCHPs. Then, a novel discovery algorithm is proposed to discover Intra-department Healthcare Process (IHP) models from medical event logs. Next, collaboration patterns among medical departments are formalized and the corresponding discovery algorithms are given. Finally, the global CCHP model is obtained by integrating all collaboration patterns and IHP models. The proposed CCHP model discovery approach is instantiated and implemented by the state-of-the-art process discovery approaches *Inductive Miner* [4] and *Split Miner* [5] to show the applicability and effectiveness.

The remainder of the paper is organized as follows. Section II presents a review of the related work. Section III reviews some terminologies and notations used throughout the paper. A cross-department collaborative healthcare process case is introduced as a motivating example in Section IV. Section V details the discovery approach for CCHP models. Section VI introduces our tool support. Section VII performs comparative evaluation. Finally, Section VIII concludes the paper.

II. Related Work

In this section, we mainly summarize the work related to healthcare process modeling and analysis, process mining and healthcare process mining.

A. Healthcare Process Modeling and Analysis

Healthcare processes can be regarded as a kind of complex discrete event processes that are subject to continuous variation over time [7]. These variations may be caused by multiple factors, e.g., different conditions of patients and multiple ways in which activities can be performed by the hospital resources (physician, nurses, and other healthcare professionals).

As a formal foundation for hospital information system modelling, a healthcare workflow model was proposed in [8] on the basis of Petri nets. This method provides a theory basis for analyzing, verifying, and optimizing HISs. In [9], a modular approach is presented for modeling healthcare systems. A case study of a public healthcare area in Zaragoza is used to demonstrate the applicability of the approach. To effectively support medical information integration and solve the interoperability problem among various HISs, a workflow model is proposed in [10] to model interactive healthcare processes. By using the Platform Independent Petri Net Editor 2 (PIPE2), its reachability graph is generated, and based on which the correctness verification approach is introduced. To support emergency response timeliness analysis, an intuitive and formal workflow modeling formalism is presented in [11] by taking task execution times into account. The applicability of the proposed approach is validated by an emergency healthcare case study. More recently, a stochastic timed Petri net based healthcare workflow simulation framework is presented to support the analysis of optimal provisioning of healthcare resources in [12]. Following this line, a stochastic timed Petri net simulation engine is presented for effective analysis of emergency health care system quality and resource provisioning in [13]. Considering privacy-requirements of the involved workflows and their mutual dependencies, a novel view-based cross-organization workflow architecture is presented to support multi-granular privacy of workflows in [14]. In [15], we have proposed a systematic modeling and verification technique for cross-department collaborative business process by extending Petri nets where the model inter-department coordination relations are supported. In addition, the soundness of a cross-department medical process is checked by performing the reachability graph analysis. Although a lot of work has been done for formal modelling and analysis of hearlthcare processes, all these work needs to create models manually from scratch. In addition, the widely accessible event data from modern hospital information systems are not fully used for automatic model creation.

B. Process Mining and Healthcare Process Mining

Process mining extracts useful information from event logs commonly generated by information systems. This technology provides new means for model discovery, process monitoring, and business improvement in various applications [2]. As one of the most well-known model discovery technique, the *Alpha Miner* [3] first defines four kinds of ordering relations. Then, a Petri net-based process model is derived from these task dependency relations. To further its application for less-structured event logs and handle the spaghetti-like models which contain all details without any hierarchies, Günther and van der Aalst proposed the fuzzy mining framework in [16]. Leemans *et al*. proposed *Inductive Miner* to discover Petri net-based process models that are guaranteed to be correct in [4]. In [17], a top-down mining approach is presented to construct the cross-organization hierarchical process models in a service out-sourcing scenario by taking as input a set of distributed heterogeneous logs. To support efficient cross-organization business process management, *Buijs et al.* presents a novel approach to compare process models and their events logs from different organizations in [18].

More recently, process mining techniques have been applied in the field of healthcare data analysis [1], e.g., the extraction of typical treatment processes from electronic medical

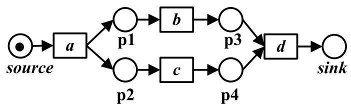


Fig. 1. Petri net example.

records [19], the automatic detection of process deviations in a trauma resuscitation [20], the discovery of role collaboration patterns among healthcare professionals [21], and the suitable abstraction of medical event data to discover processes and compare medical traces [22]. In addition, process mining approaches are used for discovering process models from medical event logs [23], checking conformance between models and event logs [24], among others. Readers may refer to [25] for a comprehensive summary.

The application of process mining techniques for analyzing medical event logs has received a lot attention in the past years. However, considering the collaboration complexity among medical departments, more efforts are required to address the problem of CCHP model discovery with high quality.

III. TERMINOLOGIES

We denote a sequence $\sigma = \langle a_1, a_2, \ldots a_n \rangle$ where $\sigma(1) = a_1, \sigma(2) = a_2, \ldots \sigma(n) = a_n$. $|\sigma| = n$ represents the length of sequence σ is n. A sequence of length 0 is called the empty sequence, denoted by $\langle\rangle$, and $\langle x \rangle$ represents a sequence with one element x. The set of all finite sequences over set S is obtained through the *Kleene star operator*, denoted by S^*. Let $\sigma \in S^*$ be a sequence, $\sigma(i)$ represents the i-th element of σ where $1 \leq i \leq |\sigma|$. Given a sequence σ and an element e, we have $e \in \sigma$ if $\exists i : 1 \leq i \leq |\sigma| \wedge \sigma(i) = e$.

In this work, we use Petri nets to represent process models and basic concepts of Petri nets are reviewed following [26].

Definition 1 (Petri Nets): A *Petri net* is defined as a 4-tuple $\Sigma = (P, T, F, M_0)$, such that

- $P \cap T = \emptyset$, $P \cup T \neq \emptyset$ where P is a finite set of places and T is a finite set of transitions;
- $F \subseteq (P \times T) \cup (T \times P)$ is a finite set of arcs; and
- $M_0 : P \rightarrow \{0, 1, 2, 3 \ldots n\}$ is the initial marking.

Places represent conditions or states and are normally drawn as small circles. Transitions represent activities or actions and are normally drawn as small rectangles. For any $x \in P \cup T$, ${}^\bullet x = \{y | (y, x) \in F\}$ is the preset of x and $x^\bullet = \{y | (x, y) \in F\}$ is its postset. M_0 denotes the initial marking and $R(M_0)$ is the set of reachable markings of Σ. For any $t \in T$, t is *enabled* under M, denoted as $(\Sigma, M)[t\rangle$, if $\forall p \in {}^\bullet t : M(p) \geq 1$. If $(\Sigma, M)[t\rangle$ holds, t may fire, resulting in a new marking M', denoted as $(\Sigma, M)[t\rangle(\Sigma, M')$ such that $M'(p) = M(p) - 1$ if $p \in {}^\bullet t \backslash t^\bullet$, $M'(p) = M(p) + 1$ if $p \in t^\bullet \backslash {}^\bullet t$, and otherwise $M'(p) = M(p)$.

An example of Petri net is shown in Fig. 1 where its initial marking is $M_0 = [source]$, and transition a is enabled under marking M_0. Transition a can fire, and the firing of a results in a new marking $M' =$ [p1, p2].

A special type of Petri net that describes a business process is called workflow net (or WF-net). Its definition is reviewed following [27].

Definition 2 (Workflow Nets): A Petri net $\Sigma = (P, T, F, M_0)$ is a WF-net if the following conditions hold:

- there exists a source place $i \in P$ such that ${}^\bullet i = \emptyset$;
- there exists a sink place $o \in P$ such that $o^\bullet = \emptyset$;
- each node $x \in P \cup T$ is on a path from i to o; and
- for each $p \in P$, $M_0(p) = 1$ if $p = i$, and otherwise $M_0(p) = 0$.

A WF-net is essentially a Petri net with a unique source place and a unique sink place. The example Petri net in Fig. 1 is a workflow net with *source* be its source place and *sink* be its sink place. A WF-net is *sound* if its corresponding short-circulated net is *1-bounded* and *live* [27].

IV. MOTIVATING EXAMPLE

In this section, a healthcare case and its corresponding log are first introduced as the motivating example, and then the medical event log is formally defined.

A. Case Description

A typical scenario of cross-department emergency medical treatment process usually involves the following departments, i.e., Surgical Department (SD), X_ray Department (XD), Cardiovascular Department (CD), Charge Office (CO), Pharmacy (PH), and Emergency Department (ED). This scenario normally includes the following steps: (1) medial staffs of ED register the patient information once emergency patients arrived; (2) medial staffs of ED perform surgery for severely injured patients, and the surgery bill is paid in the CO after surgery; (3) for lightly injured patients, medial staffs of ED perform the pre-examination, and then they enter the SD for further diagnosis and treatment; (4) according to the surgeon diagnosis, patients first go to the CO for payment, and then go to the XD to do examination; (5) after getting the imaging reports from XD, the surgeon performs diagnosis; (6) the surgeon applies for a consultation by sending consultation form to internists in the CD; (7) surgeons and internists consult to make a prescription together, and then medical staffs in the PH calculate the cost of the prescription and generate the cost bill; and (8) finally, the patient get the medicine from the PH after paying in the CO.

During the medical treatment process execution, medical event logs are recorded by the HIS. For this case, a medical event log with 18909 cases, 605088 events, and 32 medical tasks, denoted as *EM_Log*[1], is extracted from HIS. Note that this log is encrypted for privacy reasons and pre-processed to meet the *IEEE XES standard*.[2]

B. Medical Event Logs

A medical event log captured during the execution of HISs is essentially a collection of events such that each event refers to an instantiation of a medical task. In addition, a group of attributes, such as case id, timestamps, resource information, message information, and department information, are

[1] https://github.com/Lihuiling12/TASE.git
[2] https://www.xes-standard.org/

TABLE I

A FRAGMENT OF THE *EM_Log*

Event	$\#_{case}$	$\#_{task}$	$\#_{rm}$	$\#_{sm}$	$\#_{res}$	$\#_{dep}$
e_1	1	t_1	$\{m_2\}$	$\emptyset$	$\{r_1\}$	$\{SD\}$
e_2	1	t_2	$\emptyset$	$\{m_3\}$	$\emptyset$	$\{SD\}$
e_3	1	t_3	$\{m_9\}$	$\{m_4\}$	$\emptyset$	$\{SD\}$
e_4	1	t_4	$\{m_{10}\}$	$\{m_5\}$	$\emptyset$	$\{SD\}$
e_5	1	t_5	$\emptyset$	$\{m_1\}$	$\emptyset$	$\{SD\}$
e_6	1	t_6	$\{m_5\}$	$\emptyset$	$\{r_1\}$	$\{SD, CD\}$
e_7	1	t_7	$\emptyset$	$\{m_6\}$	$\emptyset$	$\{SD, CD\}$
e_8	1	t_{20}	$\{m_3\}$	$\{m_9\}$	$\emptyset$	$\{XD\}$
e_9	1	t_{21}	$\{m_4\}$	$\{m_{11}\}$	$\{r_2\}$	$\{XD\}$
e_{10}	1	t_{22}	$\{m_{12}\}$	$\emptyset$	$\emptyset$	$\{XD\}$
e_{11}	1	t_{23}	$\emptyset$	$\emptyset$	$\emptyset$	$\{XD\}$
e_{12}	1	t_{24}	$\emptyset$	$\emptyset$	$\emptyset$	$\{XD\}$
e_{13}	1	t_{25}	$\emptyset$	$\emptyset$	$\emptyset$	$\{XD\}$
e_{14}	1	t_{26}	$\emptyset$	$\{m_{10}\}$	$\emptyset$	$\{XD\}$
e_{15}	1	t_8	$\{m_1\}$	$\emptyset$	$\emptyset$	$\{CD\}$
e_{16}	1	t_9	$\emptyset$	$\emptyset$	$\emptyset$	$\{CD\}$
e_{17}	1	t_{18}	$\{m_{11}\}$	$\emptyset$	$\emptyset$	$\{CO\}$
e_{18}	1	t_{19}	$\emptyset$	$\{m_{12}\}$	$\emptyset$	$\{CO\}$
e_{19}	1	t_{16}	$\{m_7\}$	$\emptyset$	$\emptyset$	$\{CO\}$
e_{20}	1	t_{17}	$\emptyset$	$\{m_8\}$	$\emptyset$	$\{CO\}$
e_{21}	1	t_{30}	$\{m_{13}\}$	$\emptyset$	$\emptyset$	$\{CO\}$
e_{22}	1	t_{31}	$\emptyset$	$\{m_{14}\}$	$\emptyset$	$\{CO\}$
e_{23}	1	t_{13}	$\{m_6\}$	$\{m_7\}$	$\{r_2\}$	$\{PH\}$
e_{24}	1	t_{14}	$\{m_8\}$	$\emptyset$	$\emptyset$	$\{PH\}$
e_{25}	1	t_{15}	$\emptyset$	$\emptyset$	$\emptyset$	$\{PH\}$
e_{26}	1	t_{27}	$\emptyset$	$\emptyset$	$\emptyset$	$\{ED\}$
e_{27}	1	t_{10}	$\emptyset$	$\emptyset$	$\emptyset$	$\{ED\}$
e_{28}	1	t_{11}	$\emptyset$	$\emptyset$	$\emptyset$	$\{ED\}$
e_{29}	1	t_{12}	$\emptyset$	$\{m_2\}$	$\emptyset$	$\{ED\}$
e_{30}	1	t_{28}	$\emptyset$	$\{m_{13}\}$	$\{r_2\}$	$\{ED\}$
e_{31}	1	t_{29}	$\{m_{14}\}$	$\emptyset$	$\emptyset$	$\{ED\}$
e_{32}	1	t_{32}	$\emptyset$	$\emptyset$	$\emptyset$	$\{ED\}$

TABLE II

MEANING OF TASKS, MESSAGES, AND RESOURCES IN *EM_Log*

Task			
t_1	Admission	t_2	Reservation application
t_3	Imaging planning	t_4	Diagnosis
t_5	Consultation application	t_6	Consultation
t_7	Prescription given	t_8	Consultation arrangement
t_9	Consultation summary	t_{10}	Register
t_{11}	Pre-examination	t_{12}	Perform Triage
t_{13}	Medicine accounting	t_{14}	Medicine taking
t_{15}	Medicine packing	t_{16}	Payment notice of medicine
t_{17}	Payment2	t_{18}	Payment notice of X-ray
t_{19}	Payment1	t_{20}	Reservation
t_{21}	Imaging register	t_{22}	Start imaging
t_{23}	Machine operation	t_{24}	Image processing
t_{25}	Backup	t_{26}	Write report
t_{27}	Receiving patients	t_{28}	Perform Rescue
t_{29}	Pay the fees	t_{30}	Payment notice of treatment
t_{31}	Payment3	t_{32}	File storage
Message			
m_1	Consultation form	m_2	Medicine history
m_3	Reservation form	m_4	Photo form
m_5	Diagnosis report	m_6	Prescription informaiton
m_7	Cost bill2	m_8	Payment verification2
m_9	Acceptance notice	m_{10}	Imaging reports
m_{11}	Cost bill1	m_{12}	Payment verification1
m_{13}	Cost bill3	m_{14}	Payment verification3
Resource			
r_1	Diagnosis room	r_2	Charging system

involved. Formal definition of events, attributes, cases, and event logs are given as follows.

Definition 3 (Events and Attributes): Let ξ be the event universe, i.e., the set of all possible event identifiers, and N be the attribute universes, i.e., the set of all possible event attributes. For any event $e \in \xi$ and $n \in N$, $\#_n(e)$ represents the value of attribute n for e. For an arbitrary event $e \in \xi$, the following attributes are involved in medical event logs:

- $\#_{case}(e)$ represents the case to which e belongs to;
- $\#_{task}(e)$ represents the task name of e;
- $\#_{rm}(e)$ represents the set of messages received by e;
- $\#_{sm}(e)$ represents the set of messages sent by e;
- $\#_{res}(e)$ represents the set of resources required by e; and
- $\#_{dep}(e)$ represents the set of departments where e belongs to.

Definition 4 (Case, Event Log): A case over some event universe ξ is a finite sequence of events $\sigma \in \xi^*$ such that each event appears only once and all events have the same case id, i.e., $1 \leq i < j \leq |\sigma|$: $\sigma(i) \neq \sigma(j) \wedge \#_{case}(\sigma(i)) = \#_{case}(\sigma(j))$. An event log is defined as a finite set of cases, i.e., $L \subseteq \xi^*$.

Table I depicts a fragment of the medical event log *EM_Log*, and the real-life meaning of tasks, messages and resources are given in Table II. In general, this fragment involves one run of the cross-department emergency medical treatment process in Section IV-A. In total, 32 events are contained and fully ordered by the timestamp. For example, we have the following observations for e_6: (1) $\#_{case}(e_6) = 1$ means that the event belongs to case 1; (2) $\#_{task}(e_6) = t_6$ means that the task name is t_6; (3) $\#_{sm}(e_6) = \emptyset$ means that the event does not sent messages when finished; (4) $\#_{rm}(e_6) = \{m_5\}$ means that this event needs to receive message m_5 before execution; (5) $\#_{res}(e_6) = \{r_1\}$ means that this event requires resource r_1 during execution; and (6) $\#_{dep}(e_6) = \{SD, CD\}$ means that e_6 belongs to SD and CD, which indicates t_6 is a synchronous task.

V. CCHP MODEL DISCOVERY

This section first presents an approach overview. Then, the intra-department healthcare process (IHP) model discovery approach is introduced. Next, three types of collaboration patterns are formalized and discovered. Finally, a global CCHP model is obtained by integrating all discovered collaboration patterns and IHP models.

A. An Approach Overview of CCHP Model Discovery

The starting point of the proposed CCHP model discovery approach is a medical event log as introduced in Section IV-B. An approach overview is shown in Fig. 2, which includes the following four steps.

- **Medical Event Log Collection and Standardization**. During the execution of a HIS, medical event logs are recorded. To support effective model discovery, these logs should be first pre-proposed and standardized as introduced in Section IV-B.
- **IHP Model Discovery**. By taking as input a medical event log, we first project the original log to a group of sub-logs based on the department attribute. Then, a novel discovery algorithm is presented to discover an IHP model per medical department.

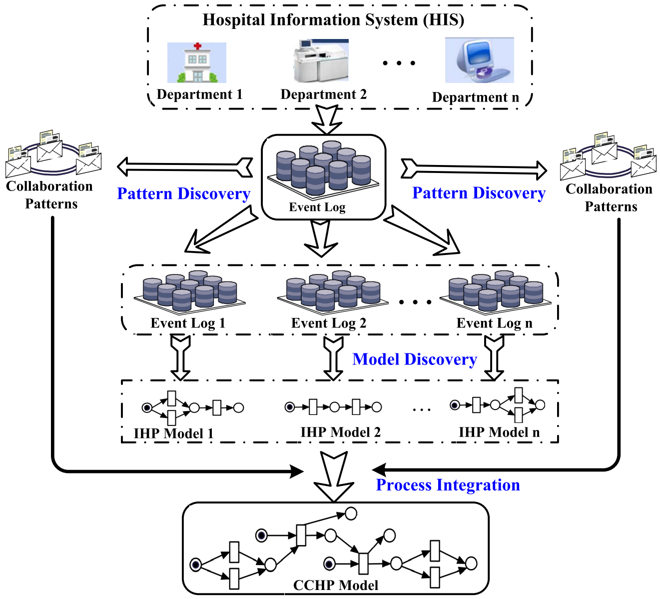


Fig. 2. Approach of overview.

- **Collaboration Pattern Discovery**. Various collaboration patterns are first formalized, and discovered by analyzing the inter-department collaboration medical event logs.
- **CCHP model discovery**. By integrating all discovered collaboration patterns and IHP models, a global CCHP model, represented as a special type of workflow net with resource and message extensions, is obtained.

B. Intra-Department Healthcare Process Model Discovery

Different from traditional business processes, the intra-department healthcare process (IHP) normally involves resource and message elements. To this end, *RM_WF_nets*, an extension of classical workflow nets with resource and message attributes, is used to formalize IHPs. The definition of *RM_WF_nets* is given following [6].

Definition 5 (RM_WF_Nets): $\Sigma = (P, T, F, M_0)$ is an RM_WF_nets such that:

- $P = P_L \cup P_R \cup P_M$ such that $P_L \cap P_R = \emptyset$, $P_L \cap P_M = \emptyset$, and $P_M \cap P_R = \emptyset$, where P_L represents the logic place set, P_M represents the message place set, and P_R represents resource place set;
- $F = F_L \cup F_R \cup F_M$ where $F_L = (P_L \times T) \cup (T \times P_L)$ represents the control-flow relation set, $F_R = (P_R \times T) \cup (T \times P_R)$ represents the resource flow relation set, and $F_M = (P_M \times T) \cup (T \times P_M)$ represents the message flow relation set; and
- $\forall p \in P$, if $p \in P_R \cup \{i\}$, $i \in P_L \wedge {}^{\bullet}i = \emptyset$, $M_0(p) = 1$, otherwise $M_0(p) = 0$.

According to [6], the resource place set (P_R) and the message place set (P_M) are separated from the normal place set P, and the firing rule of an *RM_WF_net* is the same as that of a standard WF-net.

To discover IHP model for each department, the original medical event log is projected to each department based on the department attribute. Consider the fragment of the medical event log *EM_Log* in Table I, the medical event log of the Surgical Department (SD) is obtained and shown in Table III where seven events are included.

TABLE III
FRAGMENT OF THE SURGICAL DEPARTMENT (*SD*)

Event	$\#_{case}$	$\#_{task}$	$\#_{rm}$	$\#_{sm}$	$\#_{res}$	$\#_{dep}$
e_1	1	t_1	$\{m_2\}$	$\emptyset$	$\{r_1\}$	$\{SD\}$
e_2	1	t_2	$\emptyset$	$\{m_3\}$	$\emptyset$	$\{SD\}$
e_3	1	t_3	$\{m_9\}$	$\{m_4\}$	$\emptyset$	$\{SD\}$
e_4	1	t_4	$\{m_{10}\}$	$\{m_5\}$	$\emptyset$	$\{SD\}$
e_5	1	t_5	$\emptyset$	$\{m_1\}$	$\emptyset$	$\{SD\}$
e_6	1	t_6	$\{m_5\}$	$\emptyset$	$\{r_1\}$	$\{SD\}$
e_7	1	t_7	$\emptyset$	$\{m_6\}$	$\emptyset$	$\{SD\}$

By taking as input the medical event log of single department, Algorithm 1 is presented to discover an IHP model that is represented by an *RM_WF_net*.

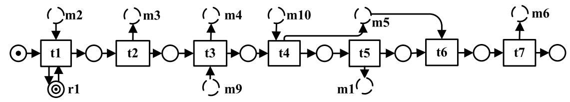
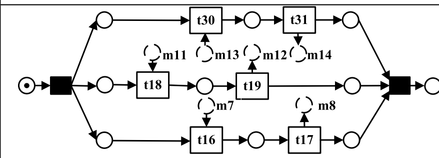
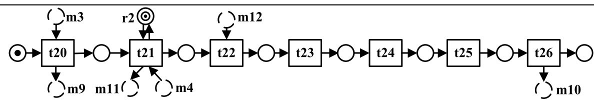
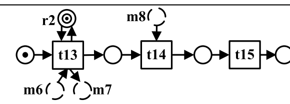
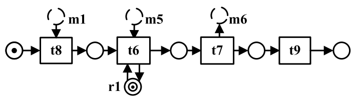
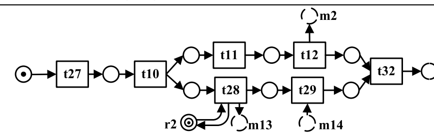
Algorithm 1 IHP Model Discovery

Input: Medical event log of a single department (L).
Output: IHP Model ($\Sigma = (P, T, F, M_0)$).
$P \leftarrow \emptyset$, $P_L \leftarrow \emptyset$, $P_R \leftarrow \emptyset$, $P_M \leftarrow \emptyset$, $F \leftarrow \emptyset$, $F_L \leftarrow \emptyset$, $F_R \leftarrow \emptyset$, $F_M \leftarrow \emptyset$, $T \leftarrow \emptyset$, $M_0 \leftarrow \emptyset$; //Initialization
$(P_L, T, F_L) \leftarrow Discovery\ Algorithm(L)$;
for $t \in T$ **do**
  for $\sigma \in L$ **do**
    for $e \in \sigma$ **do**
      if $t == \#_{task}(e)$ **then**
        if $\#_{sm}(e)! = \emptyset$ **then**
          $P_M \leftarrow P_M \cup \{\#_{sm}(e)\}$; $F_M \leftarrow F_M \cup \{(t, \#_{sm}(e))\}$;
        end if
        if $\#_{rm}(e)! = \emptyset$ **then**
          $P_M \leftarrow P_M \cup \{\#_{rm}(e)\}$; $F_M \leftarrow F_M \cup \{(\#_{rm}(e), t)\}$;
        end if
        if $\#_{res}(e)! = \emptyset$ **then**
          $P_R \leftarrow P_R \cup \{\#_{res}(e)\}$; $F_R \leftarrow F_R \cup \{(\#_{res}(e), t), (t, \#_{res}(e))\}$;
        end if
      end if
    end for
  end for
end for
$P \leftarrow P_R \cup P_L \cup P_M$, $F \leftarrow F_R \cup F_L \cup F_M$, $M_0 \leftarrow \{i\} \cup P_R$;
Return $\Sigma = (P, T, F, M_0)$.

Algorithm 1 first discovers a WF-net-based control-flow structure by applying existing discovery algorithms. To guarantee the discovered process model with high-quality, *Inductive Miner* [4] and *Split Miner* [5] are applied in this paper. Then, the event log is traversed to obtained (1) the message sending and receiving information; and (2) the resource requirement information, and finally returns an *RM_WF_net*. By taking the event log of each department as input, the obtained IHP model for each medical department is shown in Table IV.

TABLE IV

DISCOVERED IHP MODEL OF EACH MEDICAL DEPARTMENT

Medical Department	Discovered IHP Model	Medical Department	Discovered IHP Model
SD		*CO*	
XD		*PH*	
CD		*ED*	

C. Collaboration Pattern Discovery

After obtaining the IHP model per medical department, we introduce how to discover collaboration patterns among them. According to [15], the most fundamental collaboration patterns are message exchange pattern, resource sharing pattern, and task synchronization pattern in a cross-department medical process scenario. The definitions are introduced as follows.

Definition 6 (Message Exchange Pattern): Let $\Sigma_1 = (P_1, T_1, F_1, M_{01})$ and $\Sigma_2 = (P_2, T_2, F_2, M_{02})$ be the RM_WF_nets of two medical department. Message exchange patterns exist between them if: (1) $P_{L1} \cap P_{L2} = \emptyset$; (2) $P_{M1} \cap P_{M2} \neq \emptyset$; (3) $P_{R1} \cap P_{R2} = \emptyset$; and (4) $T_1 \cap T_2 = \emptyset$.

According to Definition 6, a message exchange pattern requires that there is at least one common message place between the processes of the two departments.

Definition 7 (Resource Sharing Pattern): Let $\Sigma_1 = (P_1, T_1, F_1, M_{01})$ and $\Sigma_2 = (P_2, T_2, F_2, M_{02})$ be the RM_WF_nets of the two medical departments. A resource sharing patterns exist between them if: (1) $P_{L1} \cap P_{L2} = \emptyset$; (2) $P_{M1} \cap P_{M2} = \emptyset$; (3) $P_{R1} \cap P_{R2} \neq \emptyset$; and (4) $T_1 \cap T_2 = \emptyset$.

For the resource sharing pattern, there is at least one resource shared between two departments. It is assumed that resources are exclusively used, and therefore, if a resource is occupied and locked by one medical task in one department, then other departments have to wait until it is released.

Definition 8 (Task Synchronization Pattern): Let $\Sigma_1 = (P_1, T_1, F_1, M_{01})$ and $\Sigma_2 = (P_2, T_2, F_2, M_{02})$ be the RM_WF_nets of two medical departments. Task synchronization patterns exist between them if: (1) $P_1 \cap P_2 = \emptyset$; and (2) $T_1 \cap T_2 \neq \emptyset$.

For the task synchronization pattern, there is at least one common task between the two involved departments. This common task should be undertaken by the two departments.

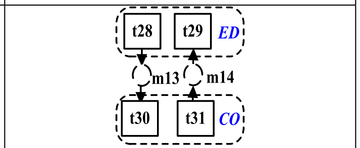
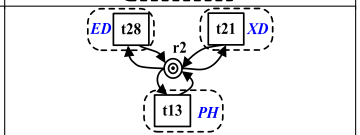
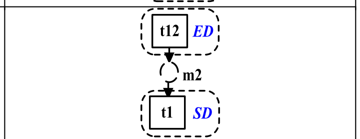
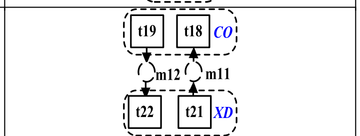
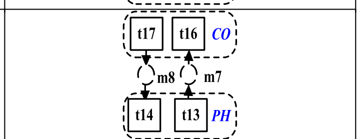
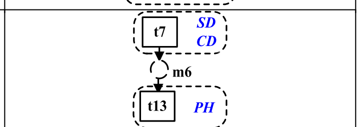
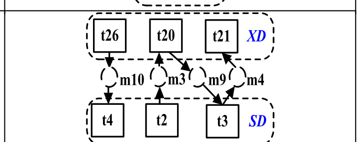
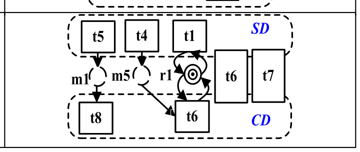
Based on Definitions 6-8, Algorithm 2 is presented to discover these three types of collaboration patterns by taking as input medical event logs. The discovered patterns are formalized by RM_WF_nets. By taking as input the medical event log *EM_Log* in Section IV-A, the discovered collaboration patterns of the cross-department emergency medical treatment process are shown in Table V by running Algorithm 2. For example, two message exchange patterns, one resource sharing pattern, and two task synchronization patterns exist between the Surgical Department (SD) and the Cardiovascular Department (CD).

Algorithm 2 Collaboration Pattern Discovery

Input: L (*Medical Event Log*).
Output: Collaboration Patterns $\Sigma = (P, T, F, M_0)$.

$P \leftarrow \emptyset$, $P_L \leftarrow \emptyset$, $P_R \leftarrow \emptyset$, $P_M \leftarrow \emptyset$, $F \leftarrow \emptyset$, $F_L \leftarrow \emptyset$, $F_R \leftarrow \emptyset$, $F_M \leftarrow \emptyset$, $T \leftarrow \emptyset$, $M_0 \leftarrow \emptyset$, $t_k \in T (k \in \{1, 2, \ldots, n\})$;
for $\sigma \in L$ **do**
  for $e_i, e_j \in \sigma$ **do**
    $t_i = \#_{task}(e_i)$; $t_j = \#_{task}(e_j)$;
    if $t_i^\bullet = {}^\bullet t_j \,\&\, {}^\bullet t_i != t_j^\bullet \,\&\, \#_{dep}(e_i) != \#_{dep}(e_j)$ **then**
      $T \leftarrow T \cup \{t_i\} \cup \{t_j\}$; $P_M \leftarrow P_M \cup \{t_i^\bullet\}$;
      $F_M \leftarrow F_M \cup \{(t_i, t_i^\bullet), (t_i^\bullet, t_j)\}$;
    end if
    if $t_i^\bullet = {}^\bullet t_j = {}^\bullet t_i = t_j^\bullet \,\&\, \#_{dep}(e_i) != \#_{dep}(e_j)$ **then**
      $T \leftarrow T \cup \{t_i\} \cup \{t_j\}$; $P_R \leftarrow P_R \cup \{t_i^\bullet\}$;
      $F_R \leftarrow F_R \cup \{(t_i, t_i^\bullet), (t_i^\bullet, t_j), (t_j, t_i^\bullet), (t_i^\bullet, t_i)\}$;
    end if
    if $|\#_{dep}(e_i)| \geq 2$ **then**
      $T \leftarrow T \cup \{t_i\}$;
    end if
  end for
end for
$P \leftarrow P_R \cup P_L \cup P_M$, $F \leftarrow F_R \cup F_L \cup F_M$, $M_0 \leftarrow P_R$;
Return $\Sigma = (P, T, F, M_0)$.
Return mined collaboration patterns.

TABLE V
DISCOVERED COLLABORATION PATTERNS

Medical Department	Discovered Collaboration Patterns
ED and CO	
ED, PH and XD	
ED and SD	
CO and XD	
CO and PH	
SD, CD and PH	
XD and SD	
SD and CD	

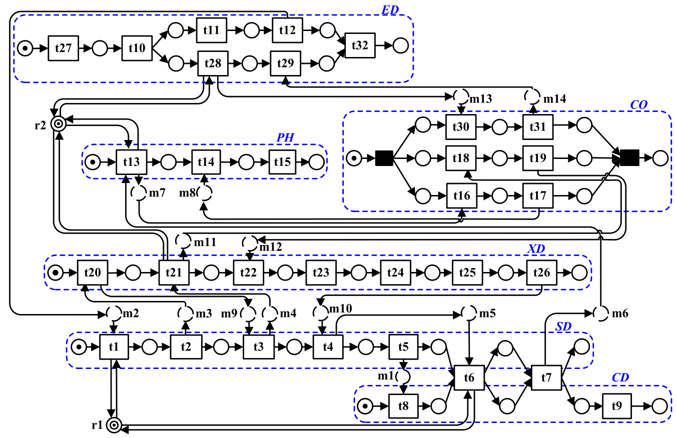


Fig. 3. Discovered CCHP model by our approach.

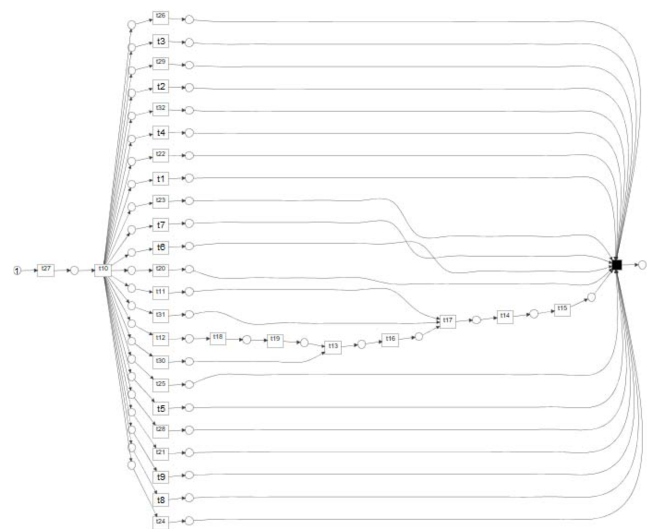


Fig. 4. Discovered model by IM.

D. Model Integration

After discovering IHP models and collaboration patterns, the global CCHP model comes into reach. In general, the CCHP model can be obtained by integrating all IHP models and collaboration patterns.

Definition 9 (Model Integration): Let $\Sigma_i = (P_i, T_i, F_i, M_{0i})(i \in \{1, \ldots n\})$ be a set of RM_WF_nets that each refers to a medical department, and $\Sigma_C = (P_C, T_C, F_C, M_{C0})$ is the set of discovered collaboration patterns. $\Sigma = (P, T, F, M_0)$ is the integrated model such that:

- $P = P_M \cup P_R \cup P_L \cup P_C$ where $P_M = P_{M1} \cup P_{M2} \cup \ldots \cup P_{Mn}$, $P_R = P_{R1} \cup P_{R2} \cup \ldots \cup P_{Rn}$, and $P_L = P_{L1} \cup P_{L2} \cup \ldots \cup P_{Ln}$;
- $T = T_C \cup T_1 \cup T_2 \cup \ldots \cup T_n$;
- $F = F_M \cup F_R \cup F_L \cup F_C$ where $F_M = F_{M1} \cup F_{M2} \cup \ldots \cup F_{Mn}$, $F_R = F_{R1} \cup F_{R2} \cup \ldots \cup F_{Rn}$, $F_L = F_{L1} \cup F_{L2} \cup \ldots \cup F_{Ln}$; and
- $M_0 = P_s \cup P_R$ where $P_s = \{p | {}^\bullet p = \emptyset, p \in P_L\}$.

According to Definition 9, the integrated model of the cross-department emergency medical treatment process is obtained and shown in Fig. 3 by taking the discovered IHP models in Table IV and the collaboration patterns in Table V as input. We have the following observations for the obtained CCHP model: (1) six medical departments are involved during the treatment process execution; and (2) these departments collaborate with each other via message exchange, resource sharing, and task synchronization. To illustrate the advantage of the proposed approach in a straightforward way, models generated by the state-of-the-art approaches, i.e., the *Inductive Miner* and *Split Miner*, by taking as input medical event logs are shown in Figs. 4-5. By comparing to Fig. 3, we see that existing discovery approaches cannot capture organization information and collaboration pattern information. In addition, discovered models are extremely complex and not precise.

According to [15], an RM_WF_net is correct if the following criteria are satisfied: (1) the workflow net of each department is sound; (2) there is no token left in message places when an *RM_WF_net* finishes; (3) for each department, the number of tokens in the sink place equals with the original number of the source place when a process finishes.

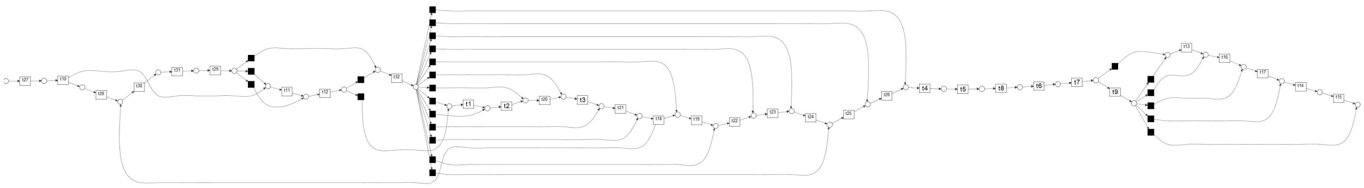


Fig. 5. Discovered model by SM.

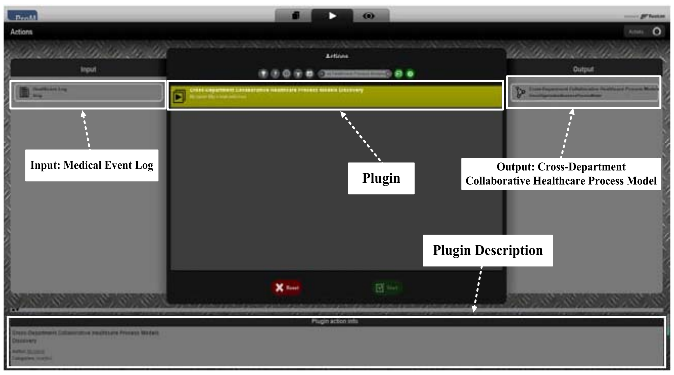


Fig. 6. Snapshot of the discovery plugin.

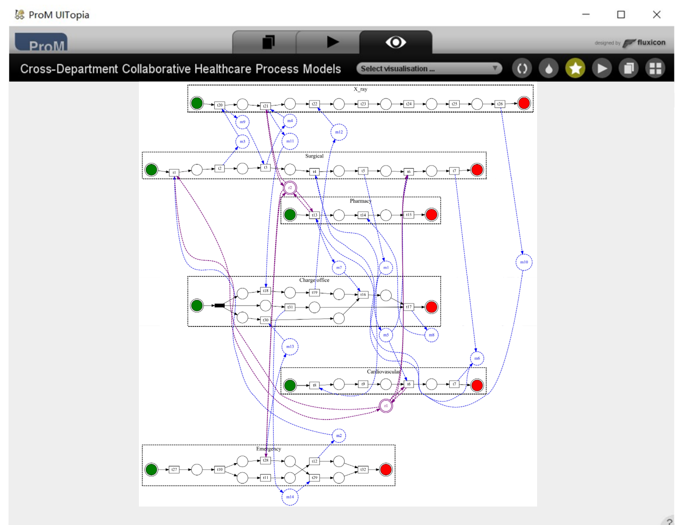


Fig. 7. Snapshot of the discovered CCHP model.

Meanwhile, there is no token in other logic places; and (4) there are no dead transitions. Note that the discovered model in Fig. 3 is correct, and therefore, it can be used for further quality evaluation by conformance checking against the input medical event log. However, the proposed approach does not guarantee that all discovered models are correct in a general case. To check the correctness of discovered CCHP models, the reachability graph analysis approach in [15] can be used.

VI. Tool Support

The open-source (Pro)cess (M)ining framework ProM 6[3] has been developed as a plugable environment for process event log analysis. The proposed CCHP model discovery approach has been implemented as a plug-in, called *Cross-Department Collaborative Healthcare Process Model Discovery*, in our ProM package[4]. This plug-in takes an XES-based medical event log as input, and returns an *RM_WF_net*. A snapshot of the tool is shown in Fig. 6.

Fig. 7 shows the discovered CCHP model represented as a *RM_WF_net* by taking as input the medical event log *EM_Log*. Note that a rectangle (a transition) represents a medical task, and a black rectangle represents a silent transition. To make the discovered model more readable, we use green place and red place to explicitly denote the start and complete of each intra-department healthcare process.

VII. Experimental Evaluation

In this section, we perform a comparative evaluation of the proposed approach against existing ones. In particular, we first introduce four public medical event logs to be used. Then, quality metrics to be used are introduced. Next, baseline discovery approaches are presented. Finally, we present the evaluation results. For these experiments, a laptop with a 2.60-GHz CPU, Windows 10, and Java SE 1.7.0 67 (64 bit) with 16 GB of allocated RAM is used.

A. Datasets

To perform the experimental evaluation, four public medical event logs[5] collected from the following scenarios, namely cross-department emergency medical treatment process event log(*EM_Log*), infectious disease medical treatment process event log (*ID_Log*), food poisoning medical treatment process event log (*FP_Log*), and special drugs management process event log (*SD_Log*), are used.

- *EM_Log*: This dataset is collected from an emergency patient treatment process in a hospital. The details of this case is given in Section IV-A;
- *ID_Log*: This dataset is collected from the treatment process of infectious diseases from a disease control and prevention center, and each trace represents the treatment process of an infectious patient;
- *FP_Log*: This dataset is collected from the response process of food poisoning incidents in a hospital, and each trace represents the response process of a food poisoning incident; and

[3]http://promtools.org/
[4]https://svn.win.tue.nl/repos/prom/Packages/ShandongPM/
[5]https://github.com/Lihuiling12/TASE.git

TABLE VI
STATISTICAL OVERVIEW OF THE EVENT LOGS

Dataset	*#Case*	*#Event*	*#Task*	*#Dep.*	*#Res.*	*#Mes.*
EM_Log	18909	605088	32	6	2	14
ID_Log	50427	1277247	30	6	0	6
FP_Log	37816	945400	27	4	0	8
SD_Log	48320	1111360	23	4	0	7

- *SD_Log*: This dataset is collected from the response process of the specially-managed drug loss in a hospital, and each trace represents the response process of a drug loss case.

Basic statistical information of these medical event logs is shown in the Table VI where #*Case* represents the number of cases, #*Event* represents the number of events, #*Task* represents the number of tasks, #*Dep*. represents the number of departments, #*Res*. represents the number of resources, and #*Mes*. represents the number of messages.

B. Quality Metrics

To evaluate the quality of discovered CCHP models against input event logs, the following quality metrics are used.

1) Fitness: Fitness quantifies the extent to which the discovered model can accurately reproduce the traces recorded in the event log. Low fitness indicates that the event log allows for much more behaviour that are not supported by the model. This section applies the fitness defined in [28].

2) Precision: Precision quantifies the fraction of the behavior allowed by the model but not allowed by the event log. Low precision means that the model allows for much more behaviour than the event log. This section applies the precision defined in [29].

Note that there is a trade off between *fitness* and *precision* [30]. Sometimes, putting aside a small amount of behaviour in the event log may cause a slight decrease in fitness while precision increases much more. Therefore, we introduce the *F-measure*. The **F-measure** is defined as the harmonic mean of *fitness* and *precision* as shown in the following equation.

$$F\text{-}measure = \frac{2 \times fitness \times precision}{fitness + precision} \tag{1}$$

C. Baseline Discovery Approaches

The following discovery approaches are compared in the experimental evaluation.

- *Inductive Miner (IM)* [4] and *Split Miner (SM)* [5] that are known as the state-of-the-art process discovery techniques are selected and denoted as **A1** and **B1**. These two approaches simply discover the CCHP model as a whole without considering the department and collaboration information.
- A divide-and-conquer framework is used by first projecting the original event log to each medical department, and then, an intra-department sub-model is discovered for each department. Finally, all discovered sub-models are merged in a parallel way. This framework is instantiated by *IM* and *SM* for intra-department discovery, denoted as **A2** and **B2**. Note that department information is supported by **A2** and **B2**, however, the collaboration patterns are not discovered.
- Our proposed CCHP model discovery approach is instantiated by *IM* and *SM* for intra-department model discovery. These two instantiations are denoted as **A3** and **B3**. Note that both the department information and collaboration patterns are discovered by **A3** and **B3**.

TABLE VII
DISCOVERY APPROACH COMPARISON

Approach	**IM**	**SM**	**Department**	**Collaboration**
A1	✓			
A2	✓		✓	
A3	✓		✓	✓
B1		✓		
B2		✓	✓	
B3		✓	✓	✓

The main features of these approaches are compared in Table VII, based on which we can see (1) **A1**-**A3** are based on *IM* and **B1**-**B3** are based on *SM*; (2) Different from **A1**, **A2** relies on the department information only and **A3** supports both department and collaboration discovery; and (3) Different from **B1**, **B2** relies on the department information only and **B3** supports both department and collaboration discovery.

D. Experimental Results

By taking the four medical event logs in Section VII-A as input, all baseline approaches are applied and evaluated in terms of *fitness*, *precision*, and *F-measure*. Detailed evaluation results are shown in Table VIII, based on which the following observations and discussions are made:

All three approaches that are build on top of the *IM*, i.e., **A1**-**A3**, can guarantee that the discovered CCHP models with perfect fitness. This observation matches the feature of *IM*. In addition, the precision of **A3** is the highest while the precision of **A1** is the lowest among these three approaches. This is because both **A2** and **A3** use the department information to localize the discovery scope within a single department, and therefore, the precision is improved compared to **A1**. Different from **A2** that simply connects all sub-models in a parallel way, **A3** that can discover collaboration behavior further improves the precision values.

As for the precision values, **B1**-**B3** that are built on the *SM* are higher than those of **A1**-**A3** that are built on *IM* for all experimental event logs. This can be explained by the fact that *IM* tries to reach a perfect fitness value by sacrificing the precision value while the *SM* aims to optimize the precision division by sacrificing the fitness division. If we focus on **B1**-**B3**, the fitness values of **B2** and **B3** are much higher than that of **B1**. This indicates that the use of department information can improve the fitness of *SM*.

Either **A3** or **B3** achieves the best F-measure values for all experimental event logs, i.e., **A3** achieves the best F-measure among all *IM*-based approaches (**A1**-**A3**), and **B3** performs the best among all *SM*-based approaches (**B1**-**B3**). Normally,

TABLE VIII
EXPERIMENTAL EVALUATION RESULTS

	A1				A2				A3			
	Fitness	Precision	F-measure	Time(ms)	Fitness	Precision	F-measure	Time(ms)	Fitness	Precision	F-measure	Time(ms)
EM_Log	**1**	0.13	0.23	4976.0	**1**	0.27	0.43	263.0	**1**	0.86	0.92	1510.0
ID_Log	**1**	0.17	0.29	7384.0	**1**	0.41	0.58	226.0	**1**	0.84	0.91	2849.0
FP_Log	**1**	0.25	0.40	7673.0	**1**	0.40	0.57	797.0	**1**	0.79	**0.88**	2627.0
SD_Log	**1**	0.32	0.48	5629.0	**1**	0.50	0.67	555.0	**1**	0.84	**0.91**	2737.0
	B1				B2				B3			
	Fitness	Precision	F-measure	Time(ms)	Fitness	Precision	F-measure	Time(ms)	Fitness	Precision	F-measure	Time(ms)
EM_Log	0.92	**0.99**	0.95	1282.0	0.98	0.31	0.47	330.0	0.98	0.97	**0.97**	1245.0
ID_Log	0.82	0.95	0.88	12050.0	**1**	0.41	0.58	493.0	0.99	**0.99**	**0.99**	2981.0
FP_Log	0.79	**0.98**	0.87	7325.0	**1**	0.40	0.57	467.0	**1**	0.79	**0.88**	2440.0
SD_Log	0.75	**0.99**	0.85	6218.0	0.99	0.49	0.66	402.0	0.99	0.71	0.83	2826.0

a high F-measure value indicates better model quality. It can be explained by the fact that **A3** and **B3** improves the quality of discovered models by explicitly supporting the cross-department collaboration behavior discovery. An exception is the *SD_Log* where the **B3** performs worse than **B1**. This can be explained by the fact that B3 improves the fitness by sacrificing the precision. Therefore, it cannot guaranteed that B3 always performs better than B1, as the increase of fitness and decrease of the precision are not deterministic.

The performance is quantified by calculating the time that a discovery approach required. Generally speaking, the less time a discovery technique spent, the more efficient it is. By taking the four experimental event logs as input, the execution time of the six discovery approaches are recorded. Note that for each event log, we run the discovery technique for five times, and the median values are recorded and shown in Table VIII. In general, **A2** achieves the best discovery efficiency among all *IM*-based techniques (i.e., **A1**, **A2**, **A3**), and **B2** achieves the best discovery efficiency among all *SM*-based techniques (i.e., **B1**, **B2**, **B3**). This is because **A2** or **B2** applies a divide-and-conquer framework by first projecting the whole log to small ones and then processing them in a parallel way. Although **A3** or **B3** are also built on top of a divide-and-conquer framework, they perform cross-department collaboration pattern discovery after discovering the intra-department processes. In this way, **A3** or **B3** is not as fast as **A2** or **B3**, but still faster than **A1** or **B1**.

As a conclusion, our proposed approach that is instantiated as **A3** and **B3** can discover CCHP models with better quality and higher efficiency compared to existing approaches.

VIII. CONCLUSION

To address the limitation of existing process model discovery techniques in the presence of cross-department collaboration behaviors, this paper proposes a novel approach to support the discovery of CCHP models from medical event logs. The presented approach has been implemented as open-source plugins in the ProM toolkit, and our experimental evaluation using publicly available medical event logs has demonstrated that the proposed approach facilitates that discovery of high-quality CCHP models, compared to the existing approaches.

As the future works, we plan to further improve and extend our approach from the following directions: (1) As the proposed approach only guarantees the correctness of all discovered intra-department processes but not the integrated CCHP models, an effective mechanism is required to enforce correctness preserving integration; (2) Considering the challenges of big event logs, we plan to apply some advanced log sampling techniques, e.g., LogRank-based sampling [31], and advanced data processing architectures, e.g., cloud computing [32], to promote the capability on handling big medical event logs; and (3) Although the proposed approach focuses on cross-department healthcare processes, it can be extended to support other collaboration scenarios, e.g., [33]–[37].

REFERENCES

[1] E. Rojas, J. Munoz-Gama, M. Sepúlveda, and D. Capurro, “Process mining in healthcare: A literature review,” *J. Biomed. Informat.*, vol. 61, pp. 224–236, Jun. 2016.

[2] W. Van Der Aalst, *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Berlin, Germany: Springer, 2011.

[3] W. van der Aalst, T. Weijters, and L. Maruster, “Workflow mining: Discovering process models from event logs,” *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 9, pp. 1128–1142, Sep. 2004.

[4] S. J. Leemans, D. Fahland, and W. van der Aalst, “Discovering block-structured process models from event logs—A constructive approach,” in *Proc. Int. Conf. Appl. Theory Petri Nets Concurrency*. Berlin, Germany: Springer, 2013, pp. 311–329.

[5] A. Augusto, R. Conforti, M. Dumas, and M. La Rosa, “Split miner: Discovering accurate and simple business process models from event logs,” in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2017, pp. 1–10.

[6] Q. Zeng, S. X. Sun, H. Duan, C. Liu, and H. Wang, “Cross-organizational collaborative workflow mining from a multi-source log,” *Decis. Support Syst.*, vol. 54, no. 3, pp. 1280–1301, 2013.

[7] P. Homayounfar, “Process mining challenges in hospital information systems,” in *Proc. Federated Conf. Comput. Sci. Inf. Syst. (FedCSIS)*, 2012, pp. 1135–1140.

[8] C. G. Yan, C. J. Jiang, Y.-Q. Shi, Z. J. Ding, and Q.-Y. Li, “Modeling and analysis for IHE workflow based on Petri net,” *J. Syst. Simulat*, vol. 18, no. 6, pp. 1696–1704, 2006.

[9] C. Mahulea, L. Mahulea, J. M. G. Soriano, and J. M. Colom, “Modular Petri net modeling of healthcare systems,” *Flexible Services Manuf. J.*, vol. 30, nos. 1–2, pp. 329–357, 2018.

[10] Z. Rong-chao, “Modeling via Petri net for IHE integration profile,” *Comput. Eng. Des.*, vol. 30, no. 23, pp. 5408–5411, 2009.

[11] J. Wang, “Emergency healthcare workflow modeling and timeliness analysis,” *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 42, no. 6, pp. 1323–1331, Nov. 2012.

[12] J. Wang, J. Tian, and R. Sun, “Emergency healthcare resource requirement analysis: A stochastic timed Petri net approach,” in *Proc. IEEE 15th Int. Conf. Netw., Sens. Control (ICNSC)*, Mar. 2018, pp. 1–6.

[13] J. Zhou, J. Wang, and J. Wang, “A simulation engine for stochastic timed Petri nets and application to emergency healthcare systems,” *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 4, pp. 969–980, Jul. 2019.

[14] K. A. Schulz and M. E. Orlowska, “Facilitating cross-organisational workflows with a workflow view approach,” *Data Knowl. Eng.*, vol. 51, no. 1, pp. 109–147, 2004.

[15] Q. Zeng, F. Lu, C. Liu, H. Duan, and C. Zhou, "Modeling and verification for cross-department collaborative business processes using extended Petri nets," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 2, pp. 349–362, Feb. 2015.

[16] C. W. Günther and W. M. Van Der Aalst, "Fuzzy mining—Adaptive process simplification based on multi-perspective metrics," in *Proc. Int. Conf. Bus. Process Manage.* Berlin, Germany: Springer, 2007, pp. 328–343.

[17] Q. Zeng, H. Duan, and C. Liu, "Top-down process mining from multi-source running logs based on refinement of Petri nets," *IEEE Access*, vol. 8, pp. 61355–61369, 2020.

[18] J. C. Buijs, B. F. van Dongen, and W. M. van der Aalst, "Towards cross-organizational process mining in collections of process models and their executions," in *Proc. Int. Conf. Bus. Process Manage.* Berlin, Germany: Springer, 2011, pp. 2–13.

[19] J. Chen, L. Sun, C. Guo, W. Wei, and Y. Xie, "A data-driven framework of typical treatment process extraction and evaluation," *J. Biomed. Informat.*, vol. 83, pp. 178–195, Jul. 2018.

[20] S. Yang *et al.*, "An approach to automatic process deviation detection in a time-critical clinical process," *J. Biomed. Informat.*, vol. 85, pp. 155–167, Sep. 2018.

[21] C. Alvarez *et al.*, "Discovering role interaction models in the emergency room using process mining," *J. Biomed. Informat.*, vol. 78, pp. 60–77, Feb. 2018.

[22] G. Leonardi, M. Striani, S. Quaglini, A. Cavallini, and S. Montani, "Leveraging semantic labels for multi-level abstraction in medical process mining and trace comparison," *J. Biomed. Informat.*, vol. 83, pp. 10–24, Jul. 2018.

[23] R. S. Mans, M. Schonenberg, M. Song, W. M. van der Aalst, and P. J. Bakker, "Application of process mining in healthcare—A case study in a Dutch hospital," in *Proc. Int. Joint Conf. Biomed. Eng. Syst. Technol.* Berlin, Germany: Springer, 2008, pp. 425–438.

[24] K. Kirchner, N. Herzberg, A. Rogge-Solti, and M. Weske, "Embedding conformance checking in a process intelligence system in hospital environments," in *Process Support and Knowledge Representation in Health Care* (Lecture Notes in Computer Science), vol. 7738. Berlin, Germany: Springer, 2012, pp. 126–139.

[25] M. Lang, T. Bürkle, S. Laumann, and H.-U. Prokosch, "Process mining for clinical workflows: Challenges and current limitations," in *MIE*, vol. 136. Amsterdam, The Netherlands: IOS Press, 2008, pp. 229–234.

[26] W. Reisig, *Petri Nets: An Introduction*. Berlin, Germany: Springer, 2012.

[27] W. M. Van Der Aalst, "The application of Petri nets to workflow management," *J. Circuits, Syst. Comput.*, vol. 8, no. 1, pp. 21–66, Feb. 1998.

[28] A. Adriansyah, B. F. van Dongen, and W. M. P. van der Aalst, "Conformance checking using cost-based fitness analysis," in *Proc. IEEE 15th Int. Enterprise Distrib. Object Comput. Conf.*, Aug. 2011, pp. 55–64.

[29] A. Adriansyah, J. Munoz-Gama, J. Carmona, B. F. van Dongen, and W. M. van der Aalst, "Alignment based precision checking," in *Proc. Int. Conf. Bus. Process Manage.* Berlin, Germany: Springer, 2012, pp. 137–149.

[30] J. De Weerdt, M. De Backer, J. Vanthienen, and B. Baesens, "A robust F-measure for evaluating discovered process models," in *Proc. IEEE Symp. Comput. Intell. Data Mining (CIDM)*, Apr. 2011, pp. 148–155.

[31] C. Liu, Y. Pei, L. Cheng, Q. Zeng, and H. Duan, "Sampling business process event logs using graph-based ranking model," *Concurrency Comput., Pract. Exper.*, vol. 33, no. 5, pp. 1–14, Mar. 2021.

[32] L. Cheng, B. F. van Dongen, and W. M. van der Aalst, "Scalable discovery of hybrid process models in a cloud computing environment," *IEEE Trans. Services Comput.*, vol. 13, no. 2, pp. 368–380, Mar. 2020.

[33] H. Duan, C. Liu, Q. Zeng, and M. Zhou, "Refinement-based hierarchical modeling and correctness verification of cross-organization collaborative emergency response processes," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 8, pp. 2548–2559, Aug. 2020.

[34] Q. Zeng, C. Liu, H. Duan, and M. Zhou, "Resource conflict checking and resolution controller design for cross-organization emergency response processes," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 10, pp. 3685–3700, Oct. 2020.

[35] C. Liu, "Hierarchical business process discovery: Identifying sub-processes using lifecycle information," in *Proc. IEEE Int. Conf. Web Services (ICWS)*, Oct. 2020, pp. 423–427.

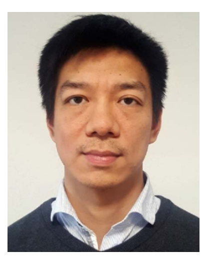
[36] C. Liu, "Formal modeling and discovery of multi-instance business processes: A cloud resource management case study," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 4, pp. 1–10, Apr. 2022.

[37] C. Liu, Q. Zeng, L. Cheng, H. Duan, M. Zhou, and J. Cheng, "Privacy-preserving behavioral correctness verification of cross-organizational workflow with task synchronization patterns," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 3, pp. 1037–1048, Jul. 2021.

Cong Liu (Member, IEEE) received the B.S. and M.S. degrees in computer software and theory from the Shandong University of Science and Technology, Qingdao, China, in 2013 and 2015, respectively, and the Ph.D. degree from the Department of Mathematics and Computer Science, Eindhoven University of Technology. He has been a Full Professor with the Shandong University of Technology since 2019. His research interests are in the areas of business process management, process mining, and Petri nets.

Huiling Li received the B.S. and M.S. degrees from the Shandong University of Technology, Zibo, China, in 2019 and 2022, respectively, where she is currently pursuing the Ph.D. degree with the School of Electrical and Electronic Engineering. Her research interests are in the areas of process mining and its applications.

Shuaipeng Zhang received the B.S. degree from Zhengzhou University, Zhengzhou, China, in 2020. He is currently pursuing the M.S. degree with the Shangdong University of Technology. His research interests are in the areas of process mining and its applications.



Long Cheng (Senior Member, IEEE) received the Ph.D. degree from the National University of Ireland in 2014. He was an Assistant Professor at Dublin City University, and a Marie Curie Fellow at University College Dublin. He is currently a Professor with the School of Control and Computer Engineering, North China Electric Power University, Beijing, and also a Visiting Professor at the Insight SFI Research Centre for Data Analytics, Dublin. His research interests include distributed systems, deep learning, cloud computing, and process mining.

Qingtian Zeng received the B.S. and M.S. degrees in computer science from the Shandong University of Science and Technology, Tai'an, China, in 1998 and 2001, respectively, and the Ph.D. degree in computer software and theory from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005. He is currently a Professor with the Shandong University of Science and Technology, Qingdao, China. His research interests are in the areas of Petri nets, process mining, and knowledge management.