

DOI:10.11918/j.issn.0367-6234.201604062

面向综合特征提取的流程相似度模型

常关羽, 杨海成, 莫蓉, 孙鹏

(现代设计与集成制造技术教育部重点实验室(西北工业大学), 西安 710072)

摘要: 针对流程相似度计算研究中注重流程结构而缺乏兼顾流程语义的问题,以及现有相似度计算方法在计算复杂度上的不足,提出一种基于流程综合特征提取的相似度计算模型. 基于流程基本控制结构分析,提出边权重标注方法以扩展现有流程结构,提取流程结构特征;定义流程高层语义模型及其对应特征提取方法;融合了节点集、边集相似度,给出新的流程结构相似度定义,利用集合关系和向量空间模型计算流程语义相似度;通过加权实现综合流程相似度评价,并采用权重参数调节的方式实现了同已有相似度计算方法的自适应转化. 将本文模型与典型相似度计算方法进行了实验对比,结果表明,面向综合特征提取的流程相似度计算方法更具普适性,同时具有更高效的计算能力.

关键词: 流程模型; 业务语义; 特征提取; 流程相似度; 流程匹配

中图分类号: TP315

文献标志码: A

文章编号: 0367-6234(2017)07-0183-06

Feature extraction oriented similarity metric of business process

CHANG Guanyu, YANG Haicheng, MO Rong, SUN Peng

(Key Laboratory of Contemporary Design and Integrated Manufacturing Technology, Ministry of Education (Northwestern Polytechnical University), Xi'an 710072, China)

Abstract: The existing research of process similarity mostly focuses on process structure but neglects the business semantic, and the similarity calculation on process structure is somehow deficient on computation complexity. To solve this problem, this paper proposes a business process representation with synthetical feature extraction, and a corresponding similarity calculation method is given at the same time. Weight notation of edges is used to extend the process structure for structure feature extraction based on the analysis of basic process control patterns, and high level business semantic is also involved for constructing business process semantic model for semantic feature extraction. The classic similarity metrics of node and edge are heuristically adapted to form a new structure similarity metric, and the similarity of business semantic is computed on vector space model and set theory. The total similarity is deduced by the weighted sum of structure and semantic similarity, and the computation model is self-adaptive to other existing methods by the adjustment of weight assignment. Finally, experiments are carried out to verify the performance of similarity computation, and the results show that the model of this paper is more adaptive and higher in computation efficiency when comparing to other methods in the literatures.

Keywords: business process model; business semantic; feature extraction; process similarity; process matching

近年来,随着业务流程管理技术和业务流程模型表达重点的转移,相关领域专家和研究学者从不同视角对流程相似度分析进行了研究^[1-2],定义了多种业务流程相似度评价模型. 作为一种对流程的有效表达方式,流程结构图及其相似度计算受到了广泛关注. 如最大共享子图匹配和图编辑距离计算都是基于图的相似度计算的代表^[3]. 基于流程的动态表现可以提取流程的动态模式,并以此模式来度量相似度. 该类方法主要特点是对流程活动节点进行标注,通过对标注信息的比较,确定流程活动相似度,即流程的相

似度是基于活动的流控结构和动态行为进行度量的^[4-5]. Wang等^[6]使用标注 PetriNet 实现业务流程模型表达,并以 PetriNet 的主要变迁序列(称为流程行为)来度量流程相似度. 而 Kunze^[5]等以流程活动配对方式提取流程特征,并基于 Jaccard 系数定义流程行为模式完成流程相似度计算. Li等^[7]在服务组合研究中,为了寻找类似的子流程,采用一种以结构为基础的相似度计算方法,通过邻接矩阵计算流程相似度. 此类方法聚焦于流程结构中单独元素,而较少考虑到流程元素间相互作用对相似度的影响. 为了提升相似度计算准确度,将与流程相关的各类资源(数据、人力、物力)加入到相似度的测定中. Baumann等^[8]考虑了与活动相关联的业务操作人、业务数据对象和活动顺序,将活动 1-1 的匹配模式改变为活动集合同 N-M 的匹配模式以计算流程相似度. Chan等^[9]

收稿日期: 2016-04-13

基金项目: 国家自然科学基金(51375395)

作者简介: 常关羽(1985—),男,博士研究生;

杨海成(1959—)男,教授,博士生导师

通信作者: 常关羽, dengxiao@mail.nwpu.edu.cn

采用类似的思路,将与活动相关的邻居活动节点信息定义为活动上下文,通过上下文比较得出流程相似度. Montani 等^[10]提出了一种知识密集性的流程相似度计算模型,该模型结合领域知识扩展了图编辑距离计算模型,将活动类型及其输入、输出纳入到流程相似度计算. 基于流程语义相似度计算主要考虑了流程节点的执行环境和业务关联,考虑更多业务背景,但业务语义较多体现在业务活动层次,对业务总体的语义体现还不充分,没有抽象出完善的业务描述体系. 虽然现有研究的成果对提升业务流程管理起到了积极推动作用,但是很少有研究能在考虑结构相似性的同时,兼顾流程较高抽象层次的语义. 另外,对于相似度度量方法的设计和评价越来越多地要求考虑其算法复杂度,以使得算法能在实际应用中发挥作用.

本文立足于建立一种较为全面的相似度计算方法,兼顾考虑流程的高层业务语义及流控结构,并通过权重分配计算综合相似度. 考虑到流程表示对流程复杂度的影响,采用特征提取的方式对流程结构进行形式化表述,在保持流程相似度计算有效性的前提下,降低计算复杂度.

1 高层语义流程模型

在业务流程系统中,业务流程是实际业务过程的体现,流程中的各个活动节点是企业内部各业务功能的表示,活动节点间通过顺序关系表达了相应的信息流向,因此流程模型主要以流程图的结构表现出来,如 PetriNet、BPMN 和 EPC 等^[8], 流程的图模型表达也成为流程建模不可或缺的一部分. 通常业务流程建模中,流程节点被分为活动节点和网关节点. 网关节点表达节点的分合流控制. 主要有 Sequence/AND-Join/AND-Split/XOR-Join/XOR-Split 等 5 种. 流程图中这 5 类结构建模成图节点时,由于其没有活动节点对应的现实意义,无法进行可识别性标注,计算相似度时造成了识别难度. 因此采用将这 5 类经典的流控类型通过对相关节点和边的权重标注的方式进行表达.

定义 1 边权重向量. 令 $e(x \rightarrow y) \in E$ 为一条流程图的有向边, x, y 分别为边 e 的两个端点. x 为起始节点, y 为终止节点. 则称 w_x 为边 e 针对 x 的权重标注. w_y 为边 e 针对 y 的权重标注. $w = (w_x, w_y)$ 为边 e 的权重向量.

根据定义 1,除 Sequence 外的 4 类典型流控的边权重向量标注见图 1.

在权重标注的流程结构中, a, b, c 为活动节点, g 为网关节点. 可以通过对 g 节点入度和出度差异判断节点的分合(Split/Join)特性. 边的权重向量则

表达了其分合的细致特性(AND/OR). 因此通过边的权重标注可充分体现流程结构特性.

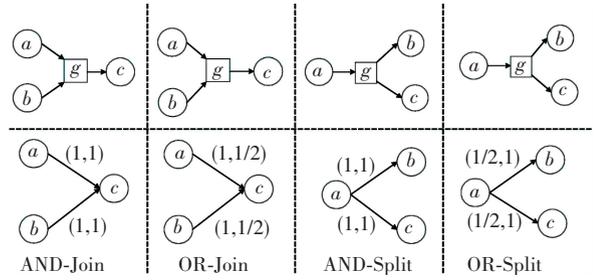


图 1 典型控制结构的权重标注

Fig.1 Weight notation of typical control pattern

定义 2 控制结构. 令 N 代表一个流程的活动节点集合,一个控制结构类型可以表达为一个三元组 $c = (x, E_x, \mu)$.

式中: $x \in N$ 为一个活动节点; $E_x \subset N \times N$ 为与该节点相关的所有边, $\forall e \in E_x, x$ 必为 e 的端点之一; $\mu: E_x \rightarrow \{(w_1, w_2) \mid \forall w_1, w_2 \in (0, 1]\}$ 为一个控制结构类型相关的边权重赋值映射.

定义 3 业务流程高层语义元数据. 业务语义的高层语义元数据可以表示为 $H = (O, A, R, G)$. 式中 O 为与业务流程相关的业务领域知识,定义了描述流程高层业务语义的必要领域概念; A 为流程的业务主角 Actor,可以是具体人,也可以是人物角色; R 为流程过程中涉及到的资源 Resource; $G = (q_1, q_2, \dots, q_i), i = 1, 2, 3, \dots \forall q_i \in Q$, 代表了业务流程的业务目标 Goal,通常以业务相关的各类业务指标 q_i 来衡量, Q 为业务指标集合.

定义 4 业务流程模型. 一个业务流程模型可以被形式化为元组 $P = (N, E, L, C, H, \lambda, \omega, \phi)$. 其中: N 为活动节点集合; $E = N \times N$ 为连接两个节点的边集合; C 为节点关联的控制结构集合; H 为业务流程的高层语义描述元数据; $\lambda: N \rightarrow L$ 为从节点到节点标签的映射; $\omega: E \rightarrow \{(w_1, w_2) \mid \forall w_1, w_2 \in (0, 1]\}$ 节点的控制结构权重赋值的映射; $\phi: N \rightarrow C$ 为将节点映射到节点关联控制结构的映射函数.

在业务流程模型中,若存在有向边 (n_1, n_2) , 则称节点 n_1 为节点 n_2 的输入,节点 n_2 为节点 n_1 的输出. 没有任何输入的节点称为开始节点,没有任何输出的节点称为结束节点.

2 流程模型特征提取及相似度计算

流程相似度较多地应用在流程资源库的检索中^[11-12],因此需要控制相似度算法的复杂度,以提升响应速度. 降低复杂度要求,简化相似度计算相关变量,应用特征提取的方法可有效简化相似度的计算变量.

2.1 结构特征提取及其相似度

定义5 流程结构特征 (S_f). $P = (N, E, L, C, H, \lambda, \omega, \phi)$ 为业务流程, 则其流程结构特征定义为 $S_f = (N, E, \omega)$. 该结构特征中 N, E, ω 具备流程定义中一致的表达意义. 容易验证这3个结构表述对象中能够确定唯一一个流程的结构. 对流程结构而言, 典型的相似度有基于节点集的相似度和基于边集的相似度.

对于流程模型 $P_1 = (N_1, E_1, L_1, C_1, H_1, \lambda_1, \omega_1, \phi_1)$ 和 $P_2 = (N_2, E_2, L_2, C_2, H_2, \lambda_2, \omega_2, \phi_2)$ 可得流程特征分别为 $S_{f1} = (N_1, E_1, \omega_1)$ 和 $S_{f2} = (N_2, E_2, \omega_2)$. 典型的节点集相似度和边集合相似度计算如下:

$$S_N = \frac{|N_1 \cap N_2|}{|N_1| + |N_2| - |N_1 \cap N_2|},$$

$$S_E = \frac{|E_1 \cap E_2|}{|E_1| + |E_2| - |E_1 \cap E_2|}.$$

针对提取出的结构特征, 对以上两种相似度计算方式进行融合, 可得基于特征提取的相似度计算方法:

$$S_{S_f} = \frac{S_N}{|N_1 \cap N_2|} \cdot \sum_{i=1}^{|N_1 \cap N_2|} S_c(\phi(n_{1i}), \phi(n_{2i})),$$

$$S_c(c_x, c_y) = \frac{1}{|E_x \cap E_y|} \sum_{j=1}^{|E_x \cap E_y|} \frac{\mu(e_{xj}) \cdot \mu(e_{yj})}{|\mu(e_{xj})| \cdot |\mu(e_{yj})|}.$$

2.2 流程的高层语义特征提取及其相似度

定义6 高层语义特征 (H_f). $P = (N, E, L, C, H, \lambda, \omega, \phi)$ 为业务流程, 具备语义 $H = (O, A, R, G)$, 则其高层语义特征定义为 $H_f = (A, R, G)$ 根据流程的高层语义元数据定义, 其中 G 为 G 的向量表示.

假定 $H_{f1} = (A_1, R_1, G_1)$ 和 $H_{f2} = (A_2, R_2, G_2)$ 为两个不同流程的高层语义特征. 则流程语义的相似度定义为

$$S_{H_f} = w_A S_A + w_R S_R + w_G S_G,$$

$$w_A + w_R + w_G = 1,$$

$$w_A, w_R, w_G \in [0, 1].$$

式中: S_A, S_R, S_G 分别为业务主角相似度、流程资源相似度和流程目标相似度, 且

$$S_A = \begin{cases} 1, & \text{if } A_1 = f_{\text{EqualActor}}(A_2), \\ 0, & \text{else;} \end{cases}$$

$$S_R = \frac{|R_1 \cap R_2|}{|R_1| + |R_2| - |R_1 \cap R_2|};$$

$$S_G = \frac{G_1 \cdot G_2}{|G_1| \cdot |G_2|}.$$

$f_{\text{EqualActor}}$ 为求取相同 A_{ctor} 的函数.

2.3 流程综合相似度

定义7 流程综合相似度. 对前文的不同流程模型 P_1 和 P_2 , 其综合相似度为

$$S_p(P_1, P_2) = \alpha S_{S_f}(S_{f1}, S_{f2}) + \beta S_{H_f}(H_{f1}, H_{f2}),$$

$$\alpha + \beta = 1,$$

$$\alpha, \beta \in [0, 1].$$

即流程的综合相似度等于结构特征相似度和语义特征相似度的加权求和. 其中 α, β 为调整参数, 通过调整参数的配比, 相似度计算方法可适用于不同的计算场景需求, 增强了其灵活性.

3 分析讨论

3.1 相似度计算过程分析

流程综合相似度的计算主要分为3个步骤: 特征提取、场景参数化和和相似度融合. 特征提取是对流程模型进行特征抽取, 完成从流程模型到流程特征的转化. 场景参数化主要涉及到相似度计算中各类参数的确定, 主要是语义相似度的参数 w_A, w_R, w_G 和综合相似度的调节参数 α, β . 具体流程见图2.

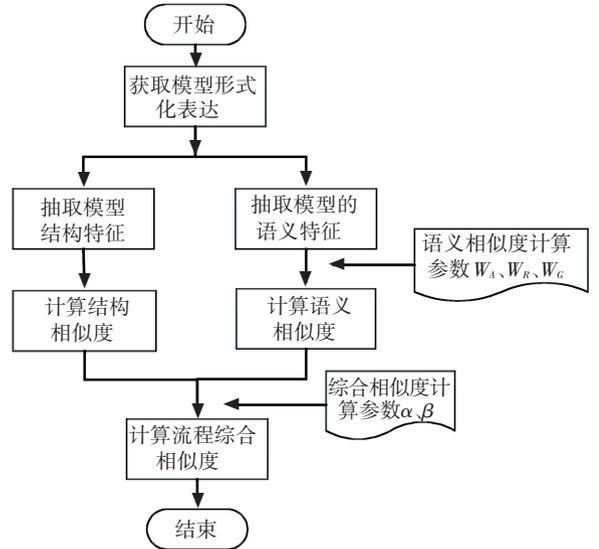


图2 相似度计算过程

Fig.2 Procedure of similarity calculation

3.2 相似度计算实例分析

1) 模型获取. 为了清楚地阐述计算过程, 对实际流程细节进行简化, 图3为两个不同的采购流程实例.

2) 形式化及特征提取. 根据对图3中的活动标识, 先对流程进行结构提取. 由定义1得到带权流程结构, 图4展示了流程的结构特征.

语义特征则从图3的业务信息中提取. 为了简化演示过程, 设定流程语义特征中 G 部分采用3维布尔向量表达 (F, T, C), 分别表示流程对 Finance/Time/Collaboration 这3类指标的要求, 对相应指标有要求则为1, 没有要求则为0. 根据语义特征的定义, 对流程 P_1 , 其涉及到对订单金额的管理, 对交付时间具备要求, 由于其完成过程涉及到的相关交互

主体很多,因此 $G_1 = (1,1,1)$, $A_1 =$ 采购单位主体,以典型的表单表达, $R_1 = \{$ 采购申请单, 采购标书, 委托交付合同, 交货清单 $\}$. 同理, 流程 P_2 主要是采

购特殊处理,对资金管理和时间交付没有固定限制,因此: $G_2 = (0,0,1)$, $A_2 =$ 采购个人主体, $R_2 = \{$ 采购申请单, 交货清单 $\}$.

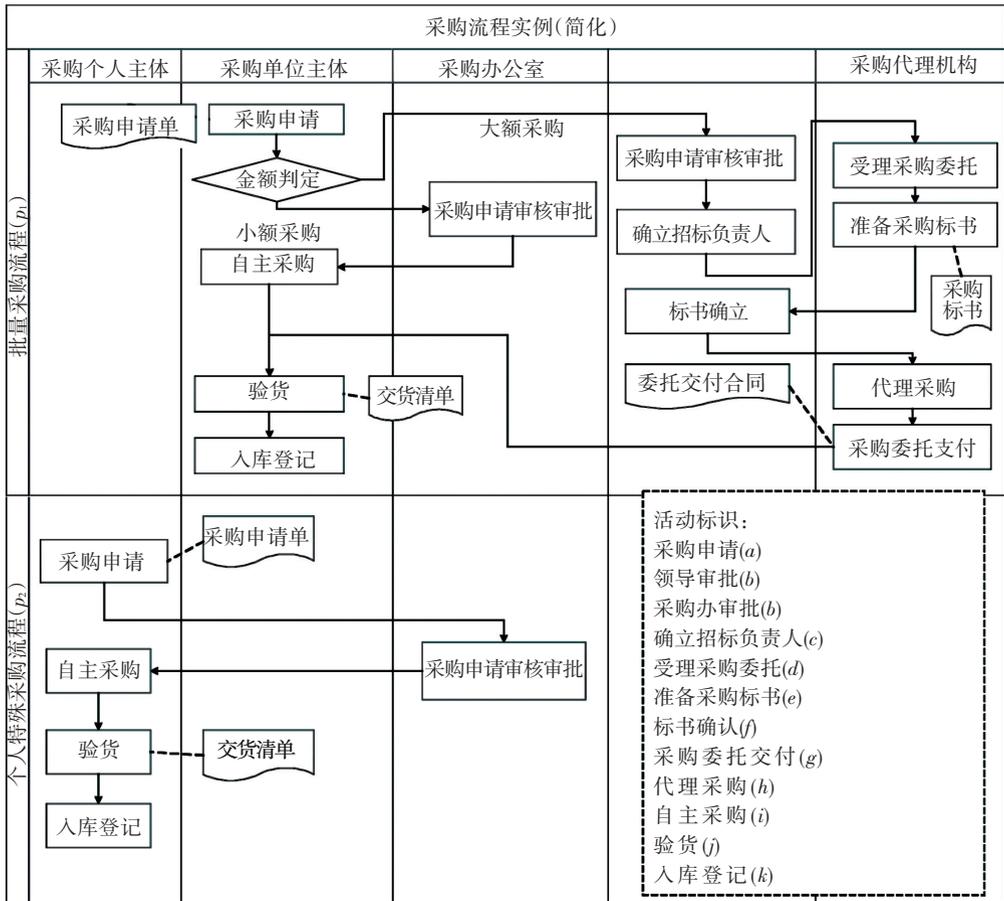


图 3 简化采购流程业务模型

Fig.3 Concise business model of procurement

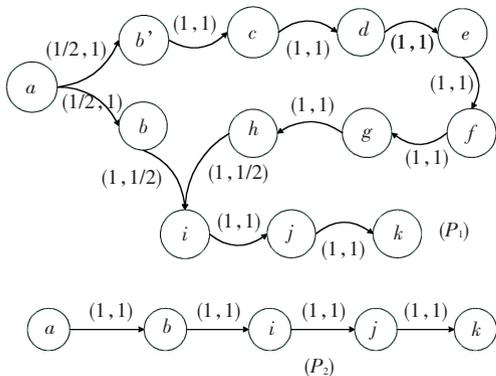


图 4 采购流程的结构形式化

Fig. 4 Formalism structure of procurement process

3) 相似度计算. 结构相似度为

$$S_{Sf} = 1/12 \times (0.3 \sqrt{10} + 0.3 \sqrt{10} + 0.5 + 0.15 \sqrt{10} + 1 + 1) = 0.406.$$

不失一般性,语义相似度计算参数采用等权重赋值,即 $w_A = w_R = w_C = 1/3$, 那么

$$S_{Hf} = w_A S_A + w_R S_R + w_C S_C = 1/3 \times 0 + 1/3 \times 2/4 + 1/3 \times 0.366 = 0.287.$$

令 $\alpha = \beta = 1/2$, 可得综合相似度

$$S_p(P_1, P_2) = 1/2 \times 0.406 + 1/2 \times 0.287 = 0.347.$$

3.3 适应性分析

相似度模型的计算准确度是衡量模型好坏的重要参考,但随着应用需求人性化和用户面的不断扩展,相似度模型的适应性也值得引起重视. 本文相似度模型,由于其设计时考虑了对流程进行较为全面的知识描述,模型信息很全面,适合除了流程建模专业人员的其他用户(业务分析人员、非 IT 系统人员等)理解和使用,相比于已有的相似度计算模型,如图编辑距离(GED)^[13]、字符串编辑距离(SED)^[14]、近距离最大子图优先(NMSF)^[15]、流程规整矩阵(PWM)^[16]等基于特征提取的相似度计算方法,其应用灵活性和用户延伸能力更强,适应性更广.

4 实验分析

实际应用中,流程的相似度主要用于流程资源库的匹配上,即用户以新的流程特征为输入,希望得到流程库中相关的流程作为输出. 实验采用 SAP

Reference 系统中提取的 600 个流程实例作为实验样本,对实验样本进行了统一的形式化和特征提取操作,形成了本文特有的兼顾高层语义的流程资源库. 实验主要从计算有效性、计算复杂度、模型参数调节效能 3 个方面考察相似度计算模型的性能. 在计算有效性方面,采用 GED、SED 和 NMSF 相似度算法进行比较. 由于该 3 种匹配算法是基于线下模式抽取和线上匹配相结合,在时间有效性上无法比较. 因此在时间有效性上采取和 NMFS 相当的 PWM 算法进行对比.

4.1 相似度计算有效性验证

因为 GED、SED、NMSF 只基于流程结构实现匹配,不支持对高层业务信息的匹配,所以在试验中采取对本文相似度模型进行纯结构匹配场景参数配置. 即令 $\alpha = 1, \beta = 0$. 在推荐结果上本文采用的方法与已有方法有所不同. 以往方法的推荐结果是点或者路径,本文推荐结果是流程库中的流程实例. 匹配的输入是通过对流程库中的流程进行特征提取后,将特征进行模糊化,并以随机化方式实现特征的部分抽取. 其他参考算法则基于各自参考文献中的伪算法描述. 实验考察了在同样的结构特征下,不同算法的匹配准确度随着匹配输出结果数的变化,流程资源库的规模为 600 个流程实例. 实验结果如图 5.

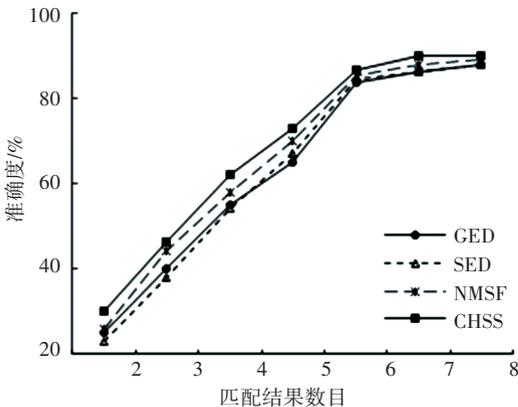


图 5 不同相似度算法准确度比较

Fig.5 Precision comparison of similarity metric

由图 5 可以看出,整体表现上,随着匹配结果数目的增加,匹配准确度不断提升,当匹配结果数目达到 6 之后,匹配准确度基本达到一个较高的稳定水平. 在匹配结果相同情况下,NMFS 算法的匹配准确度要高于 GED 和 SED 算法,这与文献中的结果吻合,证明了本文结果的可信性. 而 CHSS 算法的准确度在大部分情况下都高于其他对比算法,因此可见 CHSS 算法在相似度计算问题上的优势.

4.2 相似度计算复杂度比较

计算复杂度影响算法的处理时间,间接决定了算法在实际环境中的可应用性. 在时间复杂度比较实验

中,采用的 PWM 方法和本文算法均是线上时间衡量. 由于文献[16]中初步验证了 PWM 和 NMSF 算法时间复杂度相当,因此通过于 PWM 进行比较具备较高可信度. 实验主要考察了匹配时间随流程资源库的规模扩张导致的变化情况. 不失一般性,实验中 CHSS 采用多种参数配比最后求均值的方式进行实验,参数调控步长为 0.25. 得到的参数集合为: $(\alpha, \beta) = \{(1, 0), (0.75, 0.25), (0.5, 0.5), (0.25, 0.75), (1, 0)\}$. 对不同配比结果比较发现,同等数量的流程规模下,其耗时误差最大不超过 26 ms,属于实验正常波动范围. 不同配比多次实验数据均值见图 6.

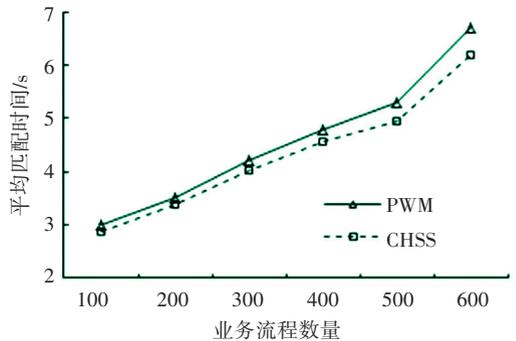


图 6 相似度算法平均匹配耗时比较

Fig.6 Average time cost of similarity calculation

图 6 为业务流程数量不断提升导致流程资源库规模扩张时,平均匹配时间的变化情况. 由图 6 可知,在不同参数配比下采用 CHSS 算法匹配的平均耗时基本同 PWM 算法相当并略有优势. 表明即使在加入语义对比的情况下,本文模型依然能保持计算高效性. 整体趋势上,资源库规模的扩张导致流程匹配的次增加,进而增加了平均匹配时间. 从平均耗时的绝对大小方面,在一般规模的流程库中应用 CHSS 算法进行匹配是能够满足应用需求的.

4.3 权重参数调节效能

综合特征提取相似度计算方法中的各类参数的可调节特性增强了其适应性,使得在结构信息不完全甚至缺乏的情况下,使用业务信息作为辅助提升模型的可用性. 在实验中主要验证在不同的 (α, β) 权重组下,相似度算法的匹配准确度表现. 实验选取 $(1, 0)$ 、 $(0, 1)$ 、 $(1/2, 1/2)$ 作为权重参数调节组合,分别命名为结构配比、语义配比和均分配比,表达了只考虑结构、只考虑语义及综合考虑两类特征对算法匹配效果的影响. 实验考察了在 3 种权重匹配下,准确度随着推荐结果数变化,结果见图 7.

由图 7 可以看出,3 种权重配比在不同匹配结果数量下对准确度均产生影响. 结构配比和语义配比在不同情况下的表现呈现波动状态,没有绝对优劣之分. 因为在随机提供流程信息进行匹配的情况

下,语义信息和结构信息的完整性是呈现波动的。因此在不同情况下,这两种配比方式会有自身的缺陷。由权重配比系数可知,均分配比的相似度输出是其余两者的算术平均。但从实验结果可知,这种均分配比的准确度也处于其余两种配比方式之间,但总是靠近表现更好的一方。因此可见,均分配比可以实现弥补其余两种匹配方式缺陷的前提下,更好地发挥其优势,实际应用中将具备更好的适应性。

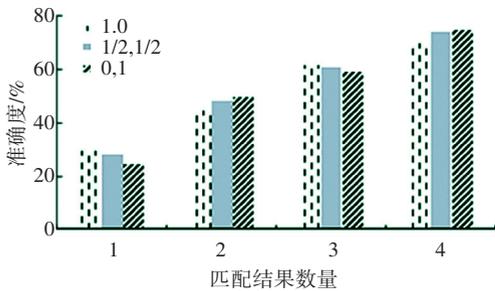


图 7 不同权重配比的准确度比较

Fig.7 Precision comparison of different weight suite

5 结 论

1) 提出一种融合高层语义的业务流程模型,在流程结构方面,通过流程控制结构定义和边权重标注对现有流程结构模型进行了扩展,根据扩展模型对典型结构相似度计算方法进行了改进。

2) 在流程语义方面,提出流程高层语义模型,阐述了其特征提取方法和基于向量空间模型的高层语义特征相似度计算方法。

3) 通过加权融合结构和语义两方面的综合特征度量流程总体相似度。并根据场景配置不同参数,提升了流程相似度计算方法的适应性。通过实验对模型进行了多方位的性能对比,结果表明:面向综合特征的相似度计算方法效率高,具备很好的应用潜力。

参 考 文 献

- [1] JIN T, WANG J, WEN L. Efficient retrieval of similar business process models based on structure [C]// On the Move to Meaningful Internet Systems. Hersonissos: Springer Berlin Heidelberg, 2011:56-63. DOI: 10.1007/978-3-642-25109-2_5.
- [2] BECKER M, LAUE R. Analysing differences between business process similarity measures [C]// BPM 2011 International Workshops. Clermont-Ferrand: Springer Berlin Heidelberg, 2012:39-49. DOI: 10.1007/978-3-642-28115-0_5.
- [3] DIJKMAN R, DUMAS M, LUCIANO G. Graph matching algorithms for business process model similarity search [C]// 7th International Conference on Business Process Management. Ulm: Springer, 2009:48-63. DOI: 10.1007/978-3-642-03848-8_5.
- [4] LIU H, LIU G, WANG Y, et al. A novel behavioral similarity measure for artifact-oriented business processes [C]// International Conference on Technology for Education and Learning. Macau: Springer Berlin Heidelberg, 2012: 81-88. DOI: 10.1007/978-3-

- 642-27711-5_12.
- [5] KUNZE M, WEIDLICH M, WESKE M. Behavioral similarity: a proper metric [C]// Business Process Management - 9th International Conference. Clermont-Ferrand: Springer Berlin Heidelberg, 2011:166-181. DOI: 10.1007/978-3-642-23059-2_15.
- [6] WANG J, HE T, WEN L, et al. A behavioral similarity measure between labeled petri-nets based on principal transition sequences [C]// On the Move Confederated International Conference. Hersonissos: Springer Berlin Heidelberg, 2010: 394-401. DOI: 10.1007/978-3-642-16934-2_27.
- [7] LI S, CAO J. A new similarity search approach on process models [C]// First International Workshop, PAS 2014. Shanghai: Springer Berlin Heidelberg, 2015:11-20. DOI: 10.1007/978-3-662-46170-9_2.
- [8] BAUMANN M H, BAUMANN M, SCHÖNIG S, et al. Towards multi-perspective process model similarity matching [C]// 10th International Workshop, EOMAS 2014. Thessaloniki: Springer Berlin Heidelberg, 2014:21-37. DOI: 10.1007/978-3-662-44860-1_2.
- [9] CHAN N N, GAALLOUL W, TATA S. Assisting business process design by activity neighborhood context matching [C]// 10th International Conference, ICSOC 2012. Shanghai: Springer Berlin Heidelberg, 2012:541-549. DOI: 10.1007/978-3-642-34321-6_38.
- [10] MONTANI S, LEONARDI G, QUAGLINI S, et al. A knowledge-intensive approach to process similarity calculation [J]. Expert Systems with Applications, 2015, 42 (9): 4207-4215. DOI: 10.1016/j.eswa.2015.01.027.
- [11] LA ROSA M, DUMAS M, EKANAYAKE C C, et al. Detecting approximate clones in business process model repositories [J]. Information Systems, 2015, 49: 102-125. DOI: 10.1016/j.is.2014.11.010.
- [12] LINCOLN M, GAL A. Searching business process repositories using operational similarity [C]// On the Move to Meaningful Internet Systems, OTM 2011. Hersonissos: Springer Berlin Heidelberg, 2011: 2-19. DOI: 10.1007/978-3-642-25109-2_2.
- [13] CAO B, YIN J, DENG S. Graph-based workflow recommendation: on improving business process modeling [C]// Proceeding of the 21st ACM International Conference on Information and Knowledge Management. New York: ACM, 2012:1527-1531. DOI:10.1145/2396761.2398466.
- [14] LI Y, CAO B, XU L, et al. An efficient recommendation method for improving business process modeling [J]. IEEE Transactions on Industrial Informatics, 2014, 10 (1): 502-513. DOI: 10.1109/THI.2013.2258677.
- [15] 曹斌, 尹建伟, 邓水光, 等. 一种基于近距离最大子图优先的业务流程推荐技术 [J]. 计算机学报, 2013, 36 (2): 263-274. DOI: 10.3724/SP.J.1016.2013.00263. CAO Bin, YIN Jianwei, DENG Shuiguang, et al. A near neighbor and maximal subgraph first based business process recommendation technique [J]. Chinese Journal of Computers, 2013, 36 (2): 263-274. DOI: 10.3724/SP.J.1016.2013.00263.
- [16] 叶岩明, 尹建伟, 曹斌. 基于流程规整矩阵的流程推荐技术 [J]. 计算机集成制造系统, 2013, 19 (8): 1868-1875. DOI: 10.13196/j.cims.2013.08.006. YE Yanming, YIN Jianwei, CAO Bin. Process warping matrix based business process recommendation technique [J]. Computer Integrated Manufacturing Systems, 2013, 19 (8): 1868-1875. DOI: 10.13196/j.cims.2013.08.006. (编辑 杨波)