

doi:10.11835/j.issn.1000-582X.2017.05.001

云制造资源语义描述和服务匹配策略

汪卫星^{1,2}

(1.重庆大学 机械传动国家重点实验室,重庆 400044;2.北海职业学院,广西 北海 536000)

摘要:针对云制造环境下资源语义描述和服务匹配问题,在充分考虑用户个性化需求与制造资源异构性的基础上,分析了云制造资源描述和服务匹配的难点,利用规范化描述的特点,将制造资源描述问题转化为Web语义描述问题,提出一种通用的制造资源语义描述框架。结合4种相似度算法,设计了阈值和权重,提出了基于过滤器的服务匹配策略。实验结果证明了该策略和算法的有效性和可行性。应用这4种相似度算法能够有效缩小匹配的规模,提高了服务的查找效率与准确度。

关键词:云制造;资源描述;匹配;相似度

中图分类号:TP391.4

文献标志码:A

文章编号:1000-582X(2017)05-001-06

Cloud manufacturing resources semantic description and service matching strategy

WANG Weixing^{1,2}

(1. College of Mechanical Engineering, Chongqing University, Chongqing 400044, P.R.China;

2. Beihai Vocational College, Beihai 536000, Guangxi, P.R.China)

Abstract: To solve the manufacturing resource description problem and the service matching problem in cloud manufacturing, a general semantic description framework for manufacturing resources which considers the user's personalized needs and the heterogeneity of manufacturing resources was proposed. The difficulties of the cloud manufacturing resource description and the service matching were analyzed. By using the features of standardized description, the manufacturing resource description problems were transformed into Web semantic description problems. Combined with four similarity algorithms and the design of threshold and weight, a service matching strategy with filter was proposed. The experimental results verified the feasibility and effectiveness of the proposed strategy and algorithm. The application of the four similarity matching algorithms can effectively reduce the matching scale and improve the service search efficiency and accuracy.

Keywords: cloud manufacturing; resource description; matching; similarity

收稿日期:2017-03-06

基金项目:国家科技支撑计划重大项目资助(2011BAF111310-02);广西高校中青年教师基础能力提升项目资助(KY2016YB755)。

Supported by the National Science and Technology Support Plan Key Projects(2011BAF111310-02) and the Basic Ability Promotion Project of Middle-Aged and Young Teachers in Colleges and Universities in Guangxi (KY2016YB755).

作者简介:汪卫星(1974-),重庆大学博士研究生,主要从事制造业信息化、云计算、网络安全等研究,(E-mail)www_ylq@126.com。

2 匹配策略

用户将制造需求通过客户端向云制造系统提出申请,云制造系统首先将用户的制造需求进行规范化描述,形成云制造资源需求本体,再通过匹配策略和算法将之与平台中已注册的云制造服务本体进行匹配,找到和用户需求相匹配的制造服务资源^[10]。匹配的理想结果应该是单项服务资源或较小范围的服务资源集,以最优化方式提供给需求者。这就需要匹配策略和算法是快捷、精准、有效、自动化或半自动化,平衡考虑查全率、查准率和计算量3个方面因素。

云制造服务资源本体化描述中的元素是众多的,其中对匹配起重要作用的、权重系数大的元素是服务资源的类型、名称信息、解释(说明、注释)、功能信息、约束信息^[11]。

1) 类型匹配:依据云制造系统中已有的类型目录进行查询,缩小匹配范围。

2) 名称匹配:依据云制造系统中的资源服务目录进行查询,得到初始匹配集。如线切割机、圆管切割机、方管切割机、异型管切割机与冷切锯通过名称匹配算法可得到它们的相似度为0,即可以直接排除,不需要进行其他属性的相似度计算。但是有时会有类似于别名的情况,如文化教育领域的“笔记本”和电子设备领域的“笔记本”是两个不同的概念,此时需要通过其他属性的相似度计算来进一步确认。所以类型匹配和解释信息匹配就起到了重要补充和修正作用^[12]。

3) 解释匹配:输入初始匹配集,对本体间的解释信息进行匹配,得到二次匹配集。

4) 功能匹配:输入二次匹配集,对本体间的功能信息进行匹配,得到三次匹配集。

5) 约束匹配:输入三次匹配集,对约束要求进行匹配,得到四次匹配集。

6) 合并匹配:将以上各种匹配集成为一个综合性的匹配过程,在此匹配过程中对以上各匹配赋予不同的权重,权重又可以根据综合匹配效果的反馈进行调整,以便优化匹配过程,使匹配结果不断优化。

3 匹配算法

3.1 类型相似度

制定规范、科学、系统的资源类型应是云制造系统得以广泛应用的重要条件,由于资源分类不规范及各种历史原因,同一名称可能表示多种异构资源,同一资源也可用多个同义名称或别名。由权威部门和行业专家共同制定行业认可的资源类型规范目录是必要的。研究采用的是按通用标准拟定的实验性质的资源类型目录,目录是分为多个层级。用户需求服务资源 R 中的类型标记为 $R_{category}$,采用目录搜索方法进行匹配,其相似度计为 $Sim_{category}(R_{category})$,返回的结果是1则表示相匹配,返回的结果是0则表示不相关。

3.2 名称相似度

名称信息主要包含服务名称描述、类型描述。使用名称信息计算相似度是本体映射过程中最直接、最基础的方法^[13]。名称信息一般是采用语言文字方式进行描述,故匹配只需要计算名称在文字上的相似度。针对语言文字类型的文本信息可采用距离方法来计算相似度。基于 WordNet 的相似度算法在概念相似度计算中得到普遍认可。它类似于一部树状语义词典,根据词义来组织词汇信息,用同义词集合(Synset)表示概念,同义词集之间是以4类明确的数学逻辑关系取得关联的,如同义与反义关系^[14]。令资源 R_1 的名称为 N_1 ,资源 R_2 的名称为 N_2 ,则 N_1 和 N_2 的语义相似度计算方法为

$$S_{name}(N_1, N_2) = \frac{2 \times d(l(N_1, N_2))}{d(N_1) + d(N_2)},$$

式中: $d(N_1)$ 和 $d(N_2)$ 分别表示名称 N_1 和名称 N_2 在WordNet树中各自的深度; $l(N_1, N_2)$ 表示2个名称的最近公共祖先。 $S_{name}(N_1, N_2)$ 取值范围在 $[0, 1]$,其值越大表示相似度越大,0表示完全不同,1表示完全相同。

为了提高搜索的准确性和效率,在此环节增加阈值(threshold value),对搜索结果进行过滤,该阈值 T_{name} 得到取值为0.55,大于阈值的匹配结果可纳入候选集合,通过匹配得到初始匹配集 $C_1\{c_1, c_2, \dots, c_i\}$,其中 i 为正整数。

3.3 解释信息相似度

云制造服务资源的解释信息是对制造资源的简介和说明,是进一步对资源名称的补充,往往是用一段文字进行描述。如果2个资源的名称及解释信息的匹配度都很高,则可初步认为这2个资源是语义相等的。资源的名称信息往往是一个名词或扩展的名称,而资源的解释信息常常是一段文本描述,所以二者的相似度匹配算法不尽相同^[15-16]。

匹配的策略的输入条件是名称匹配算法所得到的初始匹配集 $C_1\{c_1, c_2, \dots, c_i\}$ 。

首先将2个资源的解释信息进行词语解析,找出文本中的名词集合和动词集合。令资源 R_1 的解释信息为 C_1 ,解析出来 m 个名词集合为 $W_1 = \{\omega_{11}, \omega_{12}, \dots, \omega_{1m}\}$, i 个动词集合为 $V_1 = \{v_{11}, v_{12}, \dots, v_{1i}\}$ 。令资源 R_2 的解释信息为 C_2 ,解析出来 n 个名词集合为 $W_2 = \{\omega_{21}, \omega_{22}, \dots, \omega_{2n}\}$, j 个动词集合为 $V_2 = \{v_{21}, v_{22}, \dots, v_{2j}\}$ 。

3.3.1 计算名词语义相似度

先计算出 W_1, W_2 的相似度矩阵,记为

$$W_{12} = \begin{bmatrix} \omega_{11}\omega_{21} & \cdots & \omega_{1m}\omega_{21} \\ \vdots & & \vdots \\ \omega_{m1}\omega_{2n} & \cdots & \omega_{m1}\omega_{2n} \end{bmatrix},$$

矩阵中的每个元素是2个集合中单一名词对的语义相似度。遍历矩阵 W_{12} 取得每列的相似度最大值,得到该矩阵的最大相似度集合 P_L 为 $\{p_1, p_2, \dots, p_L\}$, 其中 $L = \min(m, n)$ 。

计算得到这2个解释信息的名称相似度为

$$S_{\text{noun}} = \frac{1}{l} \sum_{i=1}^l p_i,$$

式中, $1 \leq i \leq l$ 。 S_{noun} 取值范围在 $[0, 1]$, 其值越大表示相似度越大, 0 表示完全不同, 1 表示完全相同。

3.3.2 计算动词语义相似度

先计算出 V_1, V_2 的相似度矩阵,记为

$$V_{12} = \begin{bmatrix} v_{11}v_{21} & \cdots & v_{1i}v_{21} \\ \vdots & & \vdots \\ v_{i1}v_{2j} & \cdots & v_{i1}v_{2j} \end{bmatrix},$$

矩阵中的每个元素是2个集合中单一动词对的语义相似度。遍历矩阵 V_{12} 取得每列的相似度最大值,得到该矩阵的最大相似度集合 Q_L 为 $\{q_1, q_2, \dots, q_k\}$, 其中 $k = \min(i, j)$ 。

计算得到这2个解释信息的名称相似度为

$$S_{\text{verb}} = \frac{1}{k} \sum_{j=1}^k q_j,$$

式中, $1 \leq j \leq k$ 。 S_{verb} 取值范围在 $[0, 1]$, 其值越大表示相似度越大, 0 表示完全不同, 1 表示完全相同。

3.3.3 解释信息相似度计算

结合名词语义相似度、动词语义相似度,计算得到解释信息的语义相似度为

$$S_{\text{explain}} = \omega_1 S_{\text{noun}} + \omega_2 S_{\text{verb}},$$

式中: S_{noun} 为名词语义相似度, S_{verb} 为动词语义相似度, ω_1 和 ω_2 是权重, 且 $\omega_1 + \omega_2 = 1$ 。由于在解释信息中名称性语义对该服务资源的意义偏重要, 故推荐将名词语义相似度 ω_1 设为 0.55~0.75 之间的实数值。

此环节设定的阈值 T_{explain} 取值为 0.65, 大于阈值的匹配结果可纳入候选集合, 通过解释信息匹配得到二次匹配集 $C_2\{c_1, c_2, \dots, c_j\}$, 其中 j 为正整数。

3.4 集成相似度

结合名称相似度、解释信息相似度,按权重分配方法计算集成相似度,得到初选资源集合,为

$$S(N_1, N_2) = v_1 S_{\text{name}}(N_1, N_2) + v_2 S_{\text{explain}}(N_1, N_2),$$

式中: $S_{\text{name}}(N_1, N_2)$ 为名称语义相似度, $S_{\text{explain}}(N_1, N_2)$ 为解释语义相似度, 且 v_1 和 v_2 是权重, 且 $v_1 + v_2 = 1$ 。此阶段可根据资源类型来调整 v_1 和 v_2 的权重值。通过集成相似度计算得到三次匹配集 $C_3\{c_1, c_2, \dots, c_j\}$, 其中 j 为正整数。

3.5 设置过滤器

集成相似度匹配得到的结果集仍然是一个范围大,精度和准确度有待提高的中间集,需要结合其他要素对结果进一步优选^[17],缩小最终结果范围和匹配的准确度。部分学者建议进一步对制造资源服务的其他属性进行相似度计算并纳入到集成相似度计算中去,从而提高准确度,但是这种方法会大大增加系统匹配的计算量,降低系统效率,实际是不可行的。使用一些重要属性对三次匹配集 $C_3\{c_1, c_2, \dots, c_j\}$ 进行过滤,则会同时达到提高匹配结果的精度和准确率^[18]。

1)功能过滤。功能属性一般描述为 IOPE(input, output, precondition, effect),利用语义距离来计算输入(input)和输出(output)的相似度,利用本体的推理检测手段来匹配前置条件(precondition)和效果(effect),设定一个最低相似度阈值对三次匹配集 C_3 进行过滤。

2)约束过滤。制造资源的约束属性从多角度对资源服务的应用给出了限制要求,是保证制造资源能有效执行任务的基本要求。通过资源间约束的匹配来过滤结果集。

可采用并行方式将功能过滤和约束过滤结合起来构建过滤器,集成时分别设定权重进行计算,从而得到优化的结果集。

4 实验结果

根据云制造资源匹配策略和算法,笔者构建了系统模型,选取了机床加工类制造资源进行实验,拟匹配的资源集合为10组 $R_1(r_1, r_2, \dots, r_{10})$,每组样本中都有不同种类、结构相异的资源,每组样本资源数不少30个。设计服务需求集合为10组 $Q_2(q_1, q_2, \dots, q_{10})$,每个需求进行特征提取获得对应形式化描述。针对需求和资源对 $(q_1, r_1), (q_2, r_2), \dots, (q_{10}, r_{10})$ 分别实施传统关键字匹配算法和设计的综合匹配算法,实验结果见图2所示。实验中使用了 WSIL Explore 工具建立和模拟领域本体,包括类、子类和层次关系等信息。与单独应用关键字(名称)进行匹配进行比较,该算法的有效性较突出。

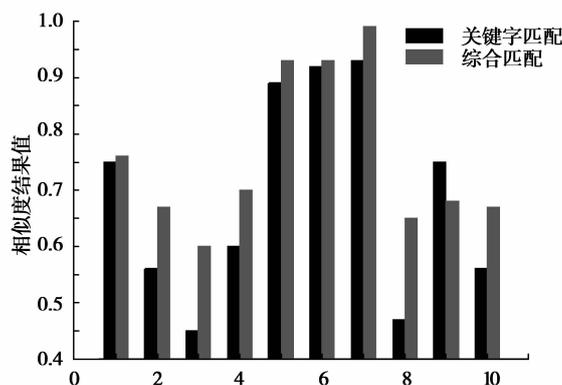


图2 2种资源匹配算法实验效果对比

Fig.2 Contrast of two resources matching algorithm experiment

5 结论

云制造资源服务化技术是云制造平台技术的重要支撑技术,采用合适信息技术对云制造资源虚拟化服务化关系到平台构建的重要因素。制造服务与制造任务之间的匹配算法与策略的优劣会极大影响平台的效率与准确度。文中对云制造资源语义描述和资源匹配策略与算法进行了研究,提出了基于 OWL-S 的通用云制造资源描述方法,给出了基本模型。结合本体语义相似度计算方法,分别对资源服务的名称、类型、注释等属性进行了相似度计算,在名称、注释等元素匹配时设置了动态阈值来增加过滤性,从而提高了匹配的效率。通过对初始结果集,应用功能和性能约束要求,进一步缩小了结果集的范围,提高了匹配的准确率。结合实验对系统的策略和算法进行了对比分析,验证了该策略和算法的科学性和实用性。今后的研究将进一步围绕以下几个方面进行:一是匹配的要素还应考虑哪些因素,二是各阶段的阈值调整到何种程度较好,三是集成匹配时各阶段结果的权重如何进行动态调整。

参考文献:

- [1] 李伯虎,张霖,王时龙.云制造——面向服务的网络化制造新模式[J].计算机集成制造系统,2010,16(1):1-7.
LI Bohu, ZHANG Lin, WANG Shilong. Cloud manufacturing: an service-oriented networked manufacturing model[J]. Computer Integrated Manufacturing Systems,2010, 16(1):1-7. (in Chinese)
- [2] 尹超,黄必清,刘飞,等.中小企业云制造服务平台共性关键技术体系[J].计算机集成制造系统,2011,17(3):495-502.
YIN Chao, HUANG Biqing, LIU Fei, et al. Common key technology system of cloud manufacturing service platform for small and medium enterprises[J]. Computer Integrated Manufacturing Systems,2011,17(3):495-502. (in Chinese)
- [3] 王成建,刘建胜,夏芳臣.云制造服务资源多层次匹配算法研究[J].机械设计与制造,2014,5(5):250-252.
WANG Chengjian, LIU Jiansheng, XIA Fangchen. Research on multi-layer matching algorithm of cloud manufacturing service resource[J]. Machinery Design and Manufacture, 2014,5(5):250-252. (in Chinese)
- [4] Bansal S, Vidal J M. Matchmaking of web services based on the DAML-S service model[C]//proceedings of the Second International Joint Conference on Autonomouous Agents and Multiagent Systems. NewYork: ACM Press, 2013:926-927.
- [5] Klein M, Bernsteina A. Toward high-precision service retrieval[J]. Internet Computing IEEE, 2004,8(11):30-36.
- [6] Yang J, Gang G. A collaboration mechanism and manufacturing resource service sharing model for service-manufacturing system[J]. International Review on Computers and Software,2012,7(2):839-844.
- [7] 肖莹莹,李伯虎,柴旭东.云制造中的制造能力服务形式化描述方法[J].系统仿真学报,2015,9(27):2096-2107.
XIAO Yingying, LI Bohu, CHAI Xudong. Research on the formalization description method of manufacturing capability service in cloud manufacturing[J]. Journal of System Simulation, 2015,9(27): 2096-2107.(in Chinese)
- [8] 李新,董朝阳.基于本体映射的云制造资源与加工任务智能匹配[J].组合机床与自动化加工技术,2015,11:157-160.
LI Xin, DONG Chaoyang. A method of intelligent matching technique of cloud manufacturing resources and processing tasks based on ontology mapping[J]. Modular Machine Tool & Automatic Manufacturing Technique,2015,11:157-160. (in Chinese)
- [9] Li H F, Jiang R, Ge S Y. Researches on manufacturing cloud service composition and optimization approach supporting for service statistic correlation[C]//Proceedings of the 26th Chinese Control and Decision Conference. Washington, D.C., USA:IEEE, 2014:4149-4154.
- [10] 康玲,陈桂松,王时龙.云制造环境下基于本体的加工资源发现[J].计算机集成制造系统,2013,19(9):2325-2331.
KANG Ling, CHENG Guisong, WANG Shilong. Ontology based process resource discovery for cloud manufacturing [J]. Computer Integrated Manufacturing Systems, 2013,19(9):2325-2331.(in Chinese)
- [11] Kessentini M, Ouni A, Langer P, et al. Search-based metamodel matching with structural and syntactic measures[J]. Journal of Systems & Software, 2014, 97:1-14.
- [12] Wang J H, Liu H, Wang H Y. A mapping-based tree similarity algorithm and its application to ontology alignment[J]. Knowledge-Based Systems, 2014, 56(C):97-107.
- [13] 易树平,刘觅,温沛涵.基于全生命周期的云制造服务研究综述[J].计算机集成制造系统,2016,4(22):871-883.
YI Shuping, LIU Mi, WEN Peihan. Overview of cloud manufacturing service based on lifecycle theory[J]. Computer Integrated Manufacturing Systems, 2016, 4(22):871-883. (in Chinese)
- [14] 赵金辉,王学慧.基于服务质量的云制造服务双向匹配模型[J].计算机集成制造系统,2016,4(22):104-112.
ZHAO Jinhui, WANG Xuehui. Two-sided matching model of cloud service based on QoS in cloud manufacturing environment[J]. Computer Integrated Manufacturing Systems, 2016, 4(22):104-112.(in Chinese)
- [15] Lartigau J, Nie L, Xu X, et al. Scheduling methodology for production services in cloud manufacturing[C] //Proceedings of International Joint Conference on Service Sciences. Washington, D.C., USA:IEEE Computer Society, 2012:34-39.
- [16] 易安斌,周宏甫.云制造环境下设备资源服务化封装方法研究[J].组合机床与自动化加工技术,2016,5:151-160.
YI Anbin, ZHOU Hongpu. Service Encapsulation method of equipment resources in cloud manufacturing environment[J]. Modular Machine Tool & Automatic Manufacturing Technique,2016,5:151-160.(in Chinese)
- [17] 周际锋,姚锡凡.面向云制造的软件资源服务化研究[J].组合机床与自动化加工技术,2016,2:157-160.
ZHOU Jifeng, YAO Xifan. Software servitization for cloud manufacturing[J]. Modular Machine Tool & Automatic Manufacturing Technique,2016,2:157-160 (in Chinese)
- [18] Yang J, Guo G. Design a new manufacturing model: cloud manufacturing [C] // Proceedings of the 2012 International Conference on Cybernetics and Informatics. Springer New York, 2013: 1597-1606.