电子商务业务流程网的可达分析方法

于汪洋[1,2,3],黄 昭[1,2],方贤文[4]
(1. 陕西师范大学现代教学技术教育部重点实验室,陕西西安 710119;2. 陕西师范大学计算机科学学院,陕西西安 710119;
3. 同济大学嵌入式系统与服务计算教育部重点实验室,上海 200092;4. 安徽理工大学理学院,安徽淮南 232001)

摘 要: 电子商务业务流程网(E-commerce Business Process Net,EBPN)是一种基于Petri网的形式化模型.该模型面向业务流程的设计阶段和应用层,整合了控制流、数据流及其相关属性,可以较好地刻画现今主流的电子业务流程,有助于描述业务流程执行过程中的数据错误和数据状态的非确定性.针对EBPN的结构和动态属性,进一步研究了EBPN的可达分析方法,给出了可达数据状态图的构造算法及相关结论.为了减少可达分析的难度,借鉴程序切片的思想,研究了EBPN的模型切片方法,定义了切片准则,构造了EBPN的切片算法.EBPN的切片方法可用于降低可达数据状态图的分析复杂度.

关键词: 电子商务;业务流程;Petri网
中图分类号: TP301 **文献标识码:** A **文章编号:** 0372-2112 (2017)07-1731-09
电子学报URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2017.07.025

Reachability Analysis Methods of E-Commerce Business Process Net

YU Wang-yang[1,2,3], HUANG Zhao[1,2], FANG Xian-wen[4]
(1. *Ministry of Education Key Laboratory for Modern Teaching Technology, Shaanxi Normal University, Xi' an, Shaanxi* 710119, *China*;
2. *School of Computer Science, Shaanxi Normal University, Xi' an, Shaanxi* 710119, *China*;
3. *The Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai* 200092, *China*;
4. *Faculty of Science, Anhui University of Science & Technology, Huainan, Anhui* 232001, *China*)

Abstract: E-commerce Business Process Nets (EBPNs) are a novel formal model for describing and validating e-commerce systems at design and application level, integrating data, control flows, and relevant attributes. Data errors and non-determinacy of the data states during the trading process can be depicted with the help of EBPNs. For static and dynamic properties of EBPN, reachability analysis methods are further studied, and construction algorithm of Reachability Data State Graph (RDSG) and related conclusions are given. In order to reduce the complexity of analyzing EBPN, referring to the program slicing, slicing of EBPN is studied. Slicing criterion and algorithm are defined and constructed. Slicing technology of EBPN can be used to reduce the complexity of analyzing EBPN.

Key words: E-commerce; business processes; petri nets

1 引言

近年来,随着网络技术的发展,以及"互联网+"相关政策的支持,网络交易作为新的商业模式发展异常迅速.然而,网络交易的安全可信问题也越发凸显.许多网络购物系统的技术不够安全和可靠,导致系统不可信.在各行业网站系统中,电子商务类网站存在高危因素比例最高,为26%[1].2014年因网络消费遭遇安全问题的网民达8000万人,占网民总数的12.6%.49.0%的网民表示互联网不太安全或非常不安全[2].因此,即使当前电子商务蓬勃发展,交易额和用户数不断攀升,但是仍然有很多用户不愿通过网络进行交易.通过技术手段确保网络交易系统的安全可信,是现代互联网发展的重要课题.网络交易遭遇不法侵害,不仅给用户带来损失,同时也严重损害了电商行业的整体信誉及网民对网络交易的信任感,严重制约了互联网经济的发展和繁荣.

为了保证网络交易系统传输的安全性,多采用安全套接层协议(SSL协议)作为底层协议.除了SSL,国

收稿日期:2015-10-08;修回日期:2016-06-10;责任编辑:孙瑶
基金项目:同济大学嵌入式系统与服务计算教育部重点实验室开放课题基金;陕西省重点科技创新团队项目(No. 2014KTC-18);国家自然科学基金(No. 61272153, No. 61402011, No. 41271387, No. 61602289);陕西省自然科学基础研究计划(No. 2016JQ6056);中央高校基本科研业务费专项资金(No. GK201503061, GK201503062)

内外学者也对各种电子商务协议做了诸多研究[3,4]. 然而,越来越多的缺陷存在于应用层和设计阶段[5,6]. 针对应用层的业务流程,国内外学者围绕业务流程的重组与优化做了很多工作[7-9]. 其中,如何确保业务流程的 Soundness 属性是研究重点[10]. 文献[11]基于 Petri 网研究了电子商务业务流程的正确性、责任及义务问题. 这些研究主要针对业务流程的设计、重构、优化或者是正确性验证,没有涉及到网络交易业务流程的行为安全问题[12-14]. 文献[15]从安全策略角度出发,研究了交互系统的互模拟问题.

流程问题是网络交易系统不可信的重要根源. 随着网络交易的不断发展,网络交易流程的实体和方式不断发生变化,开放、动态网络环境也使得网络交易系统面临的环境复杂多样. 因此,网络交易软件系统的流程设计和构造存在可信隐患会导致以业务流程为核心的网络交易系统在运行时出现不可预期的行为. 网络购物平台、第三方支付平台(TPP)及银行系统等交易主体所构成的网络交易业务流程是否完善决定着网络交易系统的可信性. 网络级和操作系统级的安全技术已经不能提供足够的保护[14,16]. 在概念模型设计阶段检测交易过程中的缺陷和逻辑错误,可以确保网络交易业务流程设计的安全性和可靠性. 如果在系统实施之后发现错误,那么对现有系统的修改和补救将是代价高昂的,很可能会造成不可挽回的损失. 针对上述问题,文献[12,13]基于 Petri 网模型,提出了基于 Petri 网的形式化模型-电子商务业务流程网(E-commerce Business Process Net,EBPN),规定了其结构属性和动态属性;在模型的概念设计阶段,针对电子商务业务流程中的逻辑缺陷,定义了交易一致性的概念,并给出了基于 EBPN 的验证技术.

本文进一步研究了 EBPN 的可达分析方法,给出了可达数据状态图的构造算法以及相关结论. 通过可达数据状态图可以分析数据有界 EBPN 的可达性、数据有界性和数据活性等性质,并且可以掌握 EBPN 所刻画的电子商务系统的所有状态,为分析电子商务系统提供了依据. 针对可达数据状态图的状态爆炸问题,研究了一种基于切片思想的可达分析方法,给出了相关定义和算法. 该方法可用于降低可达分析的难度.

2 基本概念

Petri 网是一种描述并发和分布式系统的形式化模型,并且能够刻画真并发[17]. 为了更准确地刻画电子商务系统,需要对其进行扩展. EBPN 是对原型 Petri 网的扩展,增加了数据属性、关键数据元素、关键变迁以及谓词等概念. 下面列出 EBPN 的部分基本定义和属性,更多基本概念参见文献[12,13].

定义 1[12] EBPN 是一个 7 元组 $EN=(P,T;F,D,W,S,G)$,其中:

(1)P 是有限库所集;

(2)T 是有限变迁集,且 $P\cap T=\varnothing$,$P\cup T\neq\varnothing$;

(3)$F\subseteq P\times T\cup T\times P$ 是库所与变迁之间的有向弧集;

(4)D 是 token 类型的有限集合,每一个 $d\in D$ 用一个单词来表示一个交易参数;

(5)$W:F\rightarrow <a_1d_1,a_2d_2,a_3d_3,\cdots,a_ld_l>$,$a_l\in\{0,1\}$,$d_l\in D$,$l>0$ 是 D 中所有元素的个数;

(6)S 是 token 类型的有限集合,即,关键数据元素的集合,且 $S\subset D$;

(7) $G:T\rightarrow\Pi$ 是谓词函数,其中,Π 是 D 上的布尔表达式集合.

定义 2[12] 设 $EN=(P,T;F,D,W,S,G)$ 为一个 EBPN 网,则映射 $M:P\rightarrow <n_1d_1,n_2d_2,n_3d_3,\cdots,n_ld_l>$ 为 EN 的一个标识,其中,$n_l\in\mathbb{N}$,$d_l\in D$,$l>0$ 为 D 中的元素个数.

一个标识 M 为每一个库所分配一个 l 维向量. 向量的分量 n_ld_l 表示一个库所有 n_l 个 d_l 类型的 token.

定义 3[12] 对于任意 $p\in P$,l 维向量 $M(p)$ 的多重集记为 $\widetilde{M}(p)$,而 $M(p)$ 的数据元素集合则记为 $\hat{M}(p)$. $d\in$ 在 $\widetilde{M}(p)$ 中出现的次数记为#$(d,\widetilde{M}(p))$.

为了方便表达,一个标识可以表示为 $M=[p_i(\lambda)\mid p_i$ 是拥有 token 的库所,$\lambda=\widetilde{M}(p_i)]$.

定义 4[12] 设 $EN=(P,T;F,D,W,S,G)$ 为一个 EBPN 网,二元组 $\Lambda=\langle M,\alpha\rangle$ 称之为 EN 的一个数据状态,其中,M 为 EN 的一个标识,$\alpha=(\beta,\delta_D)$ 称之为数据配置,其中 $\beta\subseteq D$,δ_D 为每一个 $d\in\beta\cup\{\hat{M}(p)\mid p\in P\}$ 分配一个 $\mathbf{T}$(true)或者 $\mathbf{F}$(false)的布尔值,使得 $\forall d\in(D-S)\rightarrow\delta_D(d)=\mathbf{T}$,$\delta_D:S\rightarrow\{\mathbf{T},\mathbf{F}\}$.

定义 5[12] 给定 $t\subset T$,$^{\cdot}t=P'$,并且 $t^{\cdot}=P''$,t 称为一个关键变迁,当且仅当:

(1) $\exists s\in\{\widetilde{W}(p,t)\mid p\in P'\}\cap S\cap\{\widetilde{W}(t,p)\mid p\in P''\}$;

(2)s 为 t 所产生的 token$\rightarrow\delta_D(s)\in\{\mathbf{T},\mathbf{F}\}$.

如果 $t\in T$ 是一个关键变迁,那么通过触发 t 产生的属于关键数据元素的 token 将会有不确定的值,即,$\mathbf{T}$(true)或者 $\mathbf{F}$(false),而这是由数据状态来体现的.

如果存在一个变迁序列 $\sigma=t_1t_2\cdots t_{k-1}$ 和数据状态序列 $<M_1,\alpha_1>$,$<M_2,\alpha_2>$,$\cdots$,$<M_k,\alpha_k>$,使得 $<M_1,\alpha_1>\xrightarrow{t_1}<M_2,\alpha_2>,\cdots,<M_{k-1},\alpha_{k-1}>\xrightarrow{t_{k-1}}<M_k,\alpha_k>$,那么称之为 $<M_k,\alpha_k>$ 从 $<M_1,\alpha_1>$ 是可达的,用表达式 $<M_1,\alpha_1>\xrightarrow{\sigma}<M_k,\alpha_k>$ 来表示. 从 $<M_1,\alpha_1>$ 可达的所有扩展数据状态,记为 $R(<M_1,\alpha_1>)$.

定义 6[12]　设 $EN=(P,T;F,D,W,S,G)$ 为一个 EBPN,$\Lambda=<M,\alpha>=<M,(\beta,\delta_D)>$ 为 EN 的一个数据状态,如果变迁 $t\in T$ 在 $<M,\alpha>$ 下是使能的,t 触发之后,新产生的标识 M'($M\xrightarrow{t}M'$)为:

$$M'(p)=\begin{cases}M(p)-W(p,t), & \text{if}\quad p\in {}^{\cdot}t-t^{\cdot}\\ M(p)+W(t,p), & \text{if}\quad p\in t^{\cdot}-{}^{\cdot}t\\ M(p)-W(p,t)+W(t,p), & \text{if}\quad p\in {}^{\cdot}t\cap t^{\cdot}\\ M(p), & \text{other}\end{cases}$$

同时,如果 t 不是关键变迁,那么产生的新的数据状态为:

$$\begin{aligned}\Lambda'&=<M',\alpha'>\\&=<M',(\beta',\delta'_D)>\\&=<M',(\beta'=\{\hat{M}(p)\mid p\in P\}\cup\beta\cup\{\widetilde{W}(t,p)\mid p\in t^{\cdot}\}\\&\quad-\{\hat{M}'(p)\mid p\in P\},\forall d\in(\beta\cup\{\hat{M}(p)\mid p\in P\})\\&\quad\rightarrow\delta'_D(d)=\delta_D(d)\wedge\forall d\in\{\widetilde{W}(t,p)\mid p\in t^{\cdot}\}\\&\quad-(\beta\cup\{\hat{M}(p)\mid p\in P\})\rightarrow\delta'_D(d)=\mathbf{T})>\end{aligned}$$

否则,如果 t 是关键变迁,那么将会产生一个数据状态集:

$$\begin{aligned}\Gamma=\{&<M',\alpha'>=<M',(\beta',\delta'_D)>\mid M\xrightarrow{t}M',\\&\beta'=\{\hat{M}(p)\mid p\in P\}\cup\beta\cup\{\widetilde{W}(t,p)\mid p\in t^{\cdot}\}\\&\quad-\{\hat{M}'(p)\mid p\in P\},\forall s\in\{\widetilde{W}(t,p)\mid p\in t^{\cdot}\}\cap S\\&\quad\rightarrow\delta'_D(s)\in\{\mathbf{T},\mathbf{F}\},\forall d\in(\beta'\cup\hat{M}'(p))\\&\quad-\{\widetilde{W}(t,p)\mid p\in t^{\cdot}\}\cap S\rightarrow\delta'_D(d)=\delta_D(d)\}\end{aligned}$$

这里,在变迁 t 触发之后,数据状态有两种变化,一个是标识的变化,一个是数据配置的变化. t 触发之后只产生一个标识,但是数据配置的变化分为两种情况,即 t 是否是一个关键变迁. 如果 t 不是一个关键变迁,只有一个数据状态产生. 否则,所产生的属于关键数据元素的 token 被分配 $\mathbf{T}$(true)或者 $\mathbf{F}$(false),每一种分配生成一个数据状态. 因此,产生一个新的数据状态集.

定义 7[12]　设 $<M_0,\alpha_0>$ 为 $EN=(P,T;F,D,W,S,G)$ 的初始数据状态. EN 为数据有界(data-bounded)的 EBPN 当且仅当 $\forall p\in P,M\in R(M_0),d\in\hat{M}(p)\rightarrow\#(d,\widetilde{M}(p))\leqslant 1$.

定义 8[12]　一个 EBPN 网,$EN=(P,T;F,D,W,S,G)$ 在初始数据状态 $<M_0,\alpha_0>$ 是可终止的当且仅当:

(1) $\exists\Delta=\{<M',\alpha'>\mid<M',\alpha'>\in R(<M_0,\alpha_0>),\forall t\in T\rightarrow\neg\ <M',\alpha'>\xrightarrow{t}\}$;

(2) $\forall<M,\alpha>\in R(<M_0,\alpha_0>),\exists\sigma=t_1,t_2,\cdots,t_k$ 使得 $<M,\alpha>\xrightarrow{\sigma}<M',\alpha'>,<M',\alpha'>\in\Delta$.

在 WF_net[8] 和原型 Petri 网中,健壮性(soundness)和无死锁性质[10,18]被用来确保工作流的结构正确性. 同样地,在电子商务业务流程中,也存在着结构合理性的要求.

定义 9[12]　一个 EBPN 网,$EN=(P,T;F,D,W,S,G)$ 在初始扩展数据状态 $<M_0,\alpha_0>$ 下是合理的当且仅当:

(1) EN 是数据有界的;

(2) EN 是可终止的;

(3) $\forall t\in T,\exists<M,\alpha>\in R(<M_0,\alpha_0>)$ 使得 $<M,\alpha>\xrightarrow{t}$.

定义 10　设 $EN=(P,T;F,D,W,S,G)$ 为一个 EBPN 网,$<M_0,\alpha_0>$ 为初始数据状态,$t\in T$. 如果对任意 $<M,\alpha>\in R(<M_0,\alpha_0>)$,都存在 $<M',\alpha'>\in R(<M,\alpha>)$,使得 $<M',\alpha'>\xrightarrow{t}$ 则称 t 是数据活的(data-live). 如果 $\forall t\in T$ 是数据活的,那么 EN 是数据活的.

定义 11[12]　一个 EBPN 网,$EN=(P,T;F,D,W,S,G)$ 在初始扩展数据状态 $<M_0,\alpha_0>$ 下是合理的,那么 EN 满足交易一致性如果 $\forall<M,\alpha>=<M,(\beta,\delta_D)>\in\Delta$ 满足下列条件:

(1) 如果 $\exists\hat{M}(p_t),\hat{M}(p_m),\hat{M}(p_s)$ 使得 TTP 处于已支付状态,商户也已处于完成交易状态,并且买方也已处于完成订单状态,那么不存在数据元素 d 使得 $d\in\beta\cup\{\hat{M}(p)\mid p\in P\}\cap S\rightarrow\delta_D(d)=\mathbf{F}$;

(2) $\exists\hat{M}(p_t)$ 使得 TTP 处于已支付状态$\leftrightarrow\exists\hat{M}(p_m)$ 使得商户处于已完成交易状态;

(3) $\exists\hat{M}(p_t)$ 使得 TTP 处于未支付状态$\leftrightarrow\exists\hat{M}(p_m)$ 使得商户处于未完成交易状态.

3　可达数据状态图及相关结论

对于数据有界的 EBPN 网,其可达数据状态集 $R(M_0,\alpha_0)$ 是一个有限集,因此参考原型 Petri 网的可达标识图的概念,提出可达数据状态图(Reachability Data State Graph,RDSG)的概念. 可达数据状态图是以 $R(<M_0,\alpha_0>)$ 为顶点集,以数据状态之间的直接可达关系为弧集构成一个有向图. 这种有向图是对现有的可达标识图的一种扩展,增加了相关的数据属性. 通过一个 EBPN 的可达数据状态图可以分析这个业务流程系统的状态变化和变迁发生序列情况,从而得知系统的有关性质和所模拟的电子商务业务流程的所有数据状态.

定义 12[12]　设 $<M_0,\alpha_0>$ 为 $EN=(P,T;F,D,W,S,G)$ 的初始数据状态. 则 EN 的可达数据状态图可以定义为一个三元组 RDSG(EN)=($B,E;L$),其中

(1) B 是顶点集,$B=R(<M_0,\alpha_0>)$;

(2) E 有向弧集,$E=\{<M_i,\alpha_i>,<M_j,\alpha_j>\mid<M_i,\alpha_i>,<M_j,\alpha_j>\in R(<M_0,\alpha_0>),\exists t_k\in T:<M_i,\alpha_i>\xrightarrow{t_k}<M_j,\alpha_j>\}$;

(3) $L:E\to T, L(<M_i,\alpha_i>,<M_j,\alpha_j>)=t_k$ 当且仅当 $<M_i,\alpha_i>\xrightarrow{t_k}<M_j,\alpha_j>$, t_k 为顶点 $<M_i,\alpha_i>$ 和 $<M_j,\alpha_j>$ 之间的有向弧的旁标. $<M_j,\alpha_j>$ 是 $<M_i,\alpha_i>$ 的后继顶点,而 $<M_i,\alpha_i>$ 是 $<M_j,\alpha_j>$ 的前继顶点.

RDSG(EN) = (B,E;L)是一个有向图,表示在初始数据状态 $<M_0,\alpha_0>$ 下,EBPN 的所有可达的数据状态. 其顶点为数据状态,弧是数据状态间的转换关系,弧上的标签为变迁. 可达数据状态图可以刻画 EBPN 模型的所有可达状态及运行情况. RDSG 扩展自 Petri 网的可达标识图,数据信息被添加至每一个标识. 因此,数据状态可以反映当前的交易状态和交易需要处理的数据信息. 通过分析 RDSG,可以验证 EBPN 的一些性质,并根据非法的可达数据状态的信息,采取适当的调整,以保证电子商务业务流程的安全性和可靠性.

对于不满足数据有界性的 EBPN,$R(M_0,\alpha_0)$是一个无限集合,不可能画出其可达数据状态图,为了用有限的形式表达一个有无限个数据状态的系统的运行情况,借鉴原型 Petri 网的可覆盖性图的概念,将趋于无限增长的某一类型的 token 的系数替换为 ω(ω 可以被看作是无穷大,对于任意一个正整数 k,使得 $\omega>k$,$\omega\leqslant\omega$,$\omega\pm k=\omega$),就构成了可覆盖性数据状态图. 可覆盖性数据状态图的相关内容已在文献[13]中阐述. 对可覆盖性数据状态图进行改造,可以构建可达数据状态图的构造算法.

图 1 是一个简单的 EBPN 示例,刻画了一个简单的电子商务业务流程,包含买方,商家和 TPP. 首先,购物者下订单,调用商家的 API 进行这项操作. 然后,订单信息,包括 gross 和 orderID,被发送回买方. 然后,根据订单信息,浏览器重定向到 TPP,付款详情记录在 TPP 中. 当付款完成后,TPP 调用商家 API 确认交易完成. 此外,使用 orderID,商家从其数据库中找到交易细节并确认. orderID 和 transactionID 是关键数据元素,用加粗单词表示;关键变迁 t_4 用双线矩形表示. 一个谓词[orderID]位于变迁 t_5 之上. 初始数据状态为 $<M_0,\alpha_0>=<[p_1(\text{TIdle}),p_2(\text{MIdle}),p_3(\text{SIdle})],\varnothing>$,表示参与交易的三方都已准备好进行交易. 图 1 是有界 EBPN,按照算法 1 构造其可达数据状态图得图 2.

算法 1 构造数据有界 EBPN 的可达数据状态图

输入:$EN=(P,T;F,D,W,S,G)$, $<M_0,\alpha_0>$

输出:RDSG(EN)

1. 以 $<M_0,\alpha_0>$ 为根结点,并标之以“New”;
2. While 存在标注为“New”的结点 Do
   - 2.1 任选一个标注为“New”的结点,设为 $<M,\alpha>$;
   - 2.2 If $\forall t\in T:\neg\ <M,\alpha>\xrightarrow{t}$
     把 $<M,\alpha>$ 标注为“Terminated node”,返回 step 2;
     End if
   - 2.3 For 每一个在 $<M,\alpha>$ 下使能的 $t\in T$ Do
     - 2.3.1 If t 不是关键变迁,生成一个新的数据状态 $<M',\alpha'>$;
       - 2.3.1.1 If $<M',\alpha'>$ 存在于 RDSG(EN)中
         从 $<M,\alpha>$ 到 $<M',\alpha'>$ 画一条有向弧,并把此弧标以 t;
         End if
       - 2.3.1.2 Else if $<M',\alpha'>$ 不存在于 RDSG(EN)
         引入一个新的结点 $<M',\alpha'>$;
         从 $<M,\alpha>$ 到 $<M',\alpha'>$ 画一条有向弧,并把此弧标以 t;
       End if
       将 $<M',\alpha'>$ 标以“New”;
       End if
     - 2.3.2 Else if t 是关键变迁
       生成一个结点集 Γ,从 $<M,\alpha>$ 到每一个 $<M',\alpha'>\in\Gamma$ 画一条有向弧,并把此弧标以 t;
       将每一个 $<M',\alpha'>\in\Gamma$ 标以“New”;
       End if
     End if
   End for
   - 2.4 移除结点 $<M,\alpha'>$ 的“New”标注;

Repeat

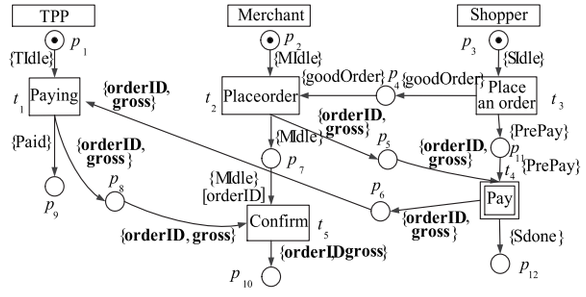

图1 一个简单的EBPN示例

算法 1 从初始数据状态 $<M_0,\alpha_0>$ 开始执行循环直到不存在未被处理的结点. 在循环的每一步,选一个数据状态 $<M,\alpha>$ 并标以“new”,并触发在 $<M,\alpha>$ 下使能的变迁. 然后,按照定义 3 产生一个数据状态或是数据状态集. 其中,关键的两个步骤是关键变迁和新产生数据状态的确定. 步骤 2.3.1 和 2.3.2 用来确定当前变迁是否为关键变迁. 步骤 2.3.1.1 确定当前新产生的数据状态是否已存在于 RDSG(EN)当中. 算法 1 针对数据有界的 EBPN 网,其可达数据状态集 $R(<M_0,\alpha_0>)$ 是一个有限集,如果当前新产生的数据状态已存在于 RDSG(EN)当中,只需要从当前的“New”结点到该结点画一条有向弧,并配置以旁标. 因此,RDSG(EN)的节点和弧的数量是有限的. 所以,算法 1 可以终止. 通过 RDSG(EN)可以分析数据有界 EBPN 的性质,并且可以分析 EBPN 的数据状态信息. 下面列出有关结论.

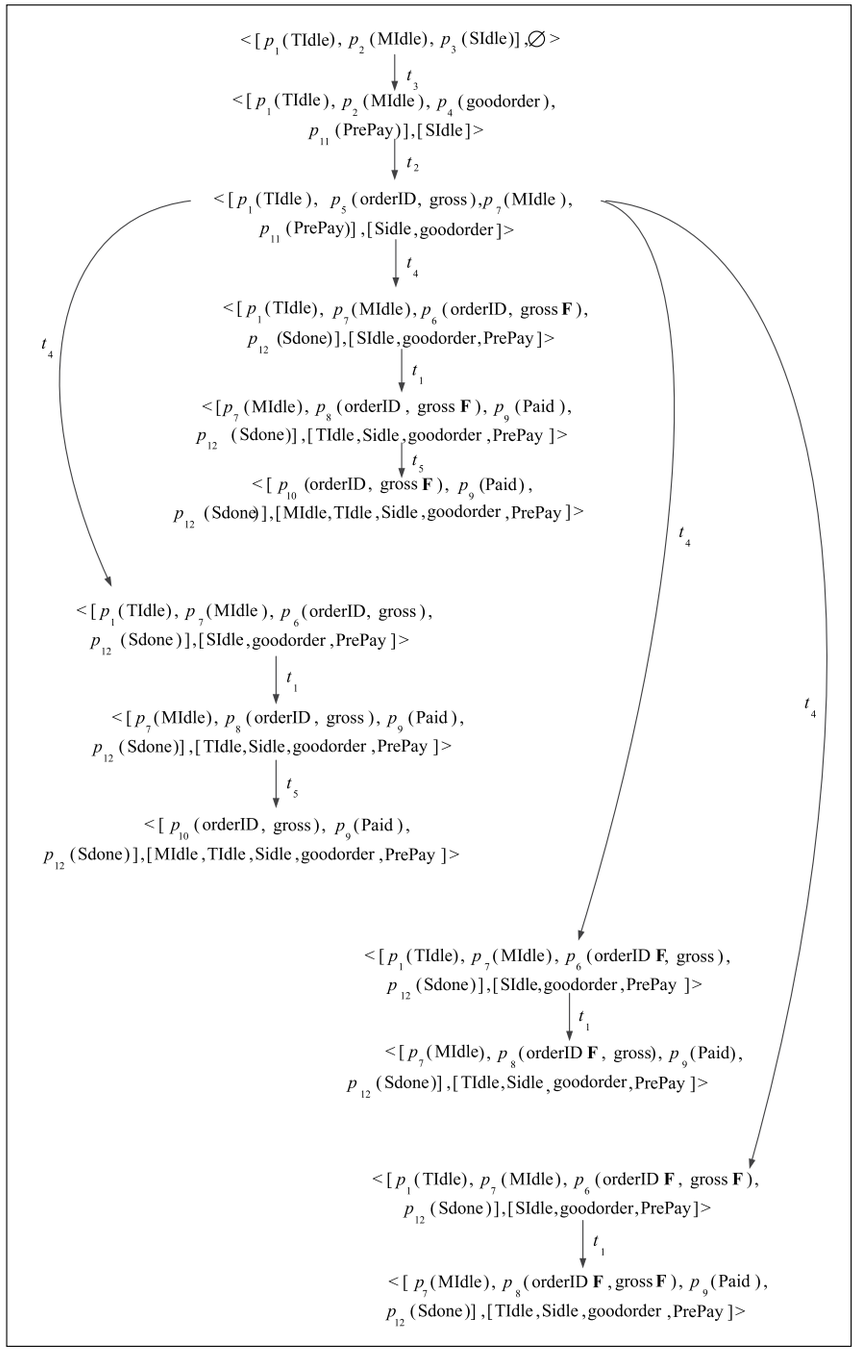

图2　可达数据状态图

引理 1　对任意 $<M,\alpha>$, $<M',\alpha'> \in R(<M_0,\alpha_0>)$, $<M',\alpha'>$ 是从 $<M,\alpha>$ 可达的当且仅当从 $<M,\alpha>$ 到 $<M',\alpha'>$ 存在一条有向路.

根据定义 6 及算法 1,可以看出该结论是显然的.

推论 1　在 RDSG(EN)中,从 $<M_0,\alpha_0>$ 到每个结点都有一条有向路.

推论 2　在 RDSG(EN)中, $<M,\alpha> \in R(<M_0,\alpha_0>)$ 是一个终端结点当且仅当 $\forall t \in T: \neg <M,\alpha> \xrightarrow{t}$.

定理 1　数据有界 EBPN 网 $EN=(P,T;F,D,W,S,G)$ 是数据活的一个充分必要条件是:在 RDSG(EN)中,从 $<M_0,\alpha_0>$ 出发的每条有向路最终都走入一个强连通子图,而且在每个这样的强连通子图中,每个 $t \in T$ 至少是一条有向弧的旁标.

证明

(1)充分性:如果在 RDSG(EN)中,每个 $t \in T$ 至少是一条有向弧的旁标,那么 $\forall t \in T$,存在 $<M', \alpha'> \in R(<M_0, \alpha_0>)$,使得 $<M', \alpha'> \xrightarrow{t}$. 因为从 $<M_0, \alpha_0>$ 出发的每条有向路最终都走入一个强连通子图,根据引理 1 和推论 1,对任意 $<M, \alpha> \in R(<M_0, \alpha_0>)$,从 $<M_0, \alpha_0>$ 到 $<M, \alpha>$ 有一条有向路,从 $<M, \alpha>$ 到 $<M', \alpha'>$ 也存在一条有向路,所以 $<M', \alpha'> \in R(<M, \alpha>)$,即对 $\forall t \in T$ 及任意 $<M, \alpha> \in R(<M_0, \alpha_0>)$,都存在 $<M', \alpha'> \in R(<M, \alpha>)$,使得 $<M', \alpha'> \xrightarrow{t}$,$EN$ 是数据活的.

(2)必要性:如果 EN 是数据活的,那么对任意 $<M, \alpha>, <M', \alpha'> \in R(<M_0, \alpha_0>)$,从 $<M_0, \alpha_0>$ 到 $<M, \alpha>, <M', \alpha'>$ 都有一条有向路,从 $<M, \alpha>$ 到 $<M', \alpha'>$ 也存在一条有向路,那么 RDSG(EN)是强连通的. 因为对任意 $<M, \alpha> \in R(<M_0, \alpha_0>)$,都存在 $<M', \alpha'> \in R(<M, \alpha>)$,使得 $<M', \alpha'> \xrightarrow{t}$,因此,每个 $t \in T$ 至少是一条有向弧的旁标.

通过分析图 2,可知 $<M_0, \alpha_0> = <[p_1(\text{TIdle}), p_2(\text{MIdle}), p_3(\text{SIdle})], \varnothing>$. 在图 2 中,从 $<M_0, \alpha_0>$ 到每个结点都有一条有向路. $<[p_{10}(\text{orderID}, \text{gross}\mathbf{F}), p_9(\text{Paid}), p_{12}(\text{Sdone})], [\text{MIdle}, \text{TIdle}, \text{SIdle}, \text{goodOrder}, \text{PrePay}]>$, $<[p_{10}(\text{orderID}, \text{gross}), p_9(\text{Paid}), p_{12}(\text{Sdone})], [\text{MIdle}, \text{TIdle}, \text{SIdle}, \text{goodOrder}, \text{PrePay}]>$, $<[p_7(\text{MIdle}), p_8(\text{orderID}\mathbf{F}, \text{gross}), p_9(\text{Paid}), p_{12}(\text{Sdone})], [\text{TIdle}, \text{SIdle}, \text{goodOrder}, \text{PrePay}]>$, $<[p_7(\text{MIdle}), p_8(\text{orderID}\mathbf{F}, \text{gross}\mathbf{F}), p_9(\text{Paid}), p_{12}(\text{Sdone})], [\text{TIdle}, \text{SIdle}, \text{goodOrder}, \text{PrePay}]>$ 为终端结点. 图 1 所示 EBPN 满足数据有界性,但不满足数据活性. 并且满足定义 8,因此也满足可终止性和结构合理性. 但是,根据定义 11,该模型不满足交易一致性. 虽然可达数据状态图可以反映电子商务业务流程所有的运行状态,但是无法避免状态爆炸问题,因而需要研究相关的有效分析方法.

4 EBPN 的切片方法

程序切片技术用于对程序进行调试、理解和维护,将比较关注的程序片段提取出来进行分析,从而减小程序分析的规模. 近年来切片思想广泛应用于形式化模型当中. 文献[19,20]研究了原型 Petri 网的切片技术,可以降低原型 Petri 网所模拟系统的分析难度. 本文借鉴原型 Petri 网的切片技术,将切片方法引入 EBPN,以避免状态爆炸带来的分析困难. 下面给出相关定义.

定义 13 设 $EN = (P, T; F, D, W, S, G)$ 为一个 EBPN 网,则 EN 的切片准则为 (Λ, Q),其中,$\Lambda = <M_0, \alpha_0> = <M_0, (\beta, \delta_D)>$ 为 EN 的初始数据状态,Q 为一个库所集,$Q \subset P$.

在分析 EBPN 模型的时候,有时只需要分析一些关键的数据状态,而不需要分析所有的状态. 在这种情况下,需要抽取出与这个状态相关的那部分模型. 切片准则是抽取模型的基准,其中,Q 表示关键数据状态所包含的库所,$\Lambda = <M_0, \alpha_0>$ 表示该 EBPN 的初始数据状态. 在切片准则中加入初始数据状态是为了考虑初始数据状态对切片抽取的影响[19].

定义 14 设 $EN = (P, T; F, D, W, S, G)$ 为一个 EBPN 网,$(\Lambda = <M_0, \alpha_0> = <M_0, (\beta, \delta_D)>, Q)$ 为的 EN 一个切片准则. 给定一个 EBPN 网 $EN' = (P', T'; F', D', W', S', G')$,$\Lambda' = <M_0', \alpha_0'> = <M_0', (\beta', \delta_D')>$ 为 EN' 的初始数据状态,其中,$M_0' = M_0 \mid_{P'}$,$\beta' \subseteq \beta$, $\forall p \in P' \subseteq P \rightarrow \delta_D'(p) = \delta_D(p)$,则 EN' 是 EN 的一个切片,下列条件必须满足:

(1)EN' 是 EN 的一个子网,即,$P' \subseteq P, T' \subseteq T, F' \subseteq F, D' \subseteq D, S' \subseteq S$, $\forall f \in F' \subseteq F, W'(f) = W(f)$; $\forall t \in T' \subseteq T, G'(t) = G(t)$;

(2)对于 EN 的初始数据状态 $\Lambda = <M_0, \alpha_0>$,如果存在一个变迁序列 $\sigma = t_1 t_2 \cdots t_n$,使得 $<M_0, \alpha_0> \xrightarrow{t_1} <M_1, \alpha_1>, \cdots, <M_{n-1}, \alpha_{n-1}> \xrightarrow{t_n} <M_n, \alpha_n>$,并且 $\exists p_i \in Q \rightarrow t_n \in {}^{\cdot} p_i$;那么在 EN' 中,在初始数据状态 $\Lambda' = <M_0', \alpha_0'>$ 下,同样存在着一个变迁序列 $\sigma = t_1 t_2 \cdots t_n$,使得 $<M_0', \alpha_0'> \xrightarrow{t_1} <M_1', \alpha_1'>, \cdots, <M_{n-1}', \alpha_{n-1}'> \xrightarrow{t_n} <M_n', \alpha_n'>$.

EN' 是 EN 的一个切片,那么 EN' 必然是 EN 的一个子网. 条件(1)描述了 EN' 是 EN 的一个子网所要满足的条件. 在影响所关注的 Q 方面,条件(2)通过对变迁行为的描述,刻画了原网和切片的行为具有等价性. 下面给出 EBPN 的切片算法,如算法 2 所示.

算法 2 构造 EBPN 的切片

输入:$EN = (P, T; F, D, W, S, G)$,$(\Lambda = <M_0, \alpha_0>, Q)$
输出:$EN' = (P', T'; F', D', W', S', G')$
1. $P' = Q, T' = \varnothing, F' = \varnothing, D' = \varnothing, W' = \varnothing, S' = \varnothing, G' = \varnothing$;
2. $P_I \subseteq P$ 为在 $\Lambda = <M_0, \alpha_0>$ 下拥有 token 的库所集合;
3. While ${}^{\cdot} Q \neq \varnothing$ Do
3.1 For 每一个 $p \in Q$ Do
$T' = T' \cup {}^{\cdot} p; P' = P' \cup {}^{\cdot} T'$;
End for
3.2 $Q = P' - Q - P_I$;
Repeat
4. $F' = F(P', T')$;
5. $D' = \{\widetilde{W}(p', t') \cup (\widetilde{W}(t'', p'') \mid p', p'' \in P'; t', t'' \in T'$;

$p' \in {}^{\cdot}t'; t'' \in {}^{\cdot}p''\}$;

6. $W' = \{W(p',t') \cup W(t'',p'') \mid p',p'' \in P'; t',t'' \in T';$
$p' \in {}^{\cdot}t'; t'' \in {}^{\cdot}p''\}$;
7. $S' = D' \cap S$;
8. $G' = \{G(t') \mid t' \in T'\}$

根据算法 2 的过程可知,算法 2 将与 Q 相关的,能够影响到 Q 的那部分子网提取出来,舍掉了不能影响 Q 的分支和结构. 算法 2 从切片准则当中的 Q 开始进行后向遍历,在 $\Lambda = <M_0, \alpha_0>$ 下,将能够往 Q 中传送 token 的子网提取出来. For 循环 3.1 返回每一个 $p \in Q$ 的前集变迁以及这些变迁的前集库所,并将其纳入到切片当中. 步骤 3.2 用于更新当前的 Q,将该次 While 循环的所有 $p \in Q$ 删除,将当前新得到库所加入. 注意算法 2 的后向遍历过程在到达初始数据状态下拥有 token 的库所时,会停止,这是考虑到初始数据状态是对 Q 影响的最初状态. 步骤 4,5,6,7,8 是获取 EN' 的其他组成部分. 因为 EBPN 中各种元素的个数是有限的,算法 2 的后向遍历过程受到步骤 3.2 的限制,直到初始数据状态算法终止. 算法复杂度基于所遍历的 EBPN 的元素个数,为 $O(|P| + |T|)$.

图 3 为文献[12]中的一个较为完整的 EBPN 模型,用于验证一个实际的网络交易业务流程. 在这个例子中{orderID,gross}为关键数据元素,一个谓词位于 t_5,用于在交易完成阶段校验关键数据元素. 开始 p_1,p_4 和 p_8 分别拥有一个类型为 TListen,MListen 和 SIdle 的 token,用来表示买方、商家和 TPP 已经准备好一次交易. 初始数据状态为 $<M_0, \alpha_0> = <[p_1(\text{TListen}), p_4(\text{MListen}), p_3(\text{SIdle})], \varnothing>$. 当 p_3,p_7 和 p_{12} 都拥有 token 的时候,表示交易已完成. t_{10} 和 t_{11} 用于刻画业务执行过程中的异常事件. 如果异常发生,则相关主体回滚至交易开始的状态.

为了分析该 EBPN 模型的交易一致性,文献[12]构造了完整的可达数据状态图,有 56 个数据状态(如图 4 所示,具体的数据状态信息参见文献[12]). 在分析过程中,如果只关注关键数据元素{orderID,gross}及其相关状态,就可以利用切片的方法来减少分析复杂度. 在这个例子中,假设关注第三方支付平台(TPP)交易完成后的状态,即,t_2 触发,p_3 和 p_{19} 获取 token 以表示支付状态,那么就可以定义一个切片准则 $(\Lambda, Q) = (<M_0, \alpha_0>, Q) = (<[p_1(\text{TListen}), p_4(\text{MListen}), p_3(\text{SIdle})], \varnothing>, \{p_3, p_{19}\})$. 基于该切片准则,利用算法 2 可得图 3 所示 EBPN 的切片(如图 5 所示). 该切片是在初始数据状态 $<M_0, \alpha_0>$ 下,能够影响 p_3 和 p_{19} 的那一部分子网. 之后,运行该切片,可得切片的可达数据状态图(如图 6 所示),仅有 23 个可达数据状态. 在分析某一个具体交易状态的时候,切片方法可以用于降低可达分析的难度.

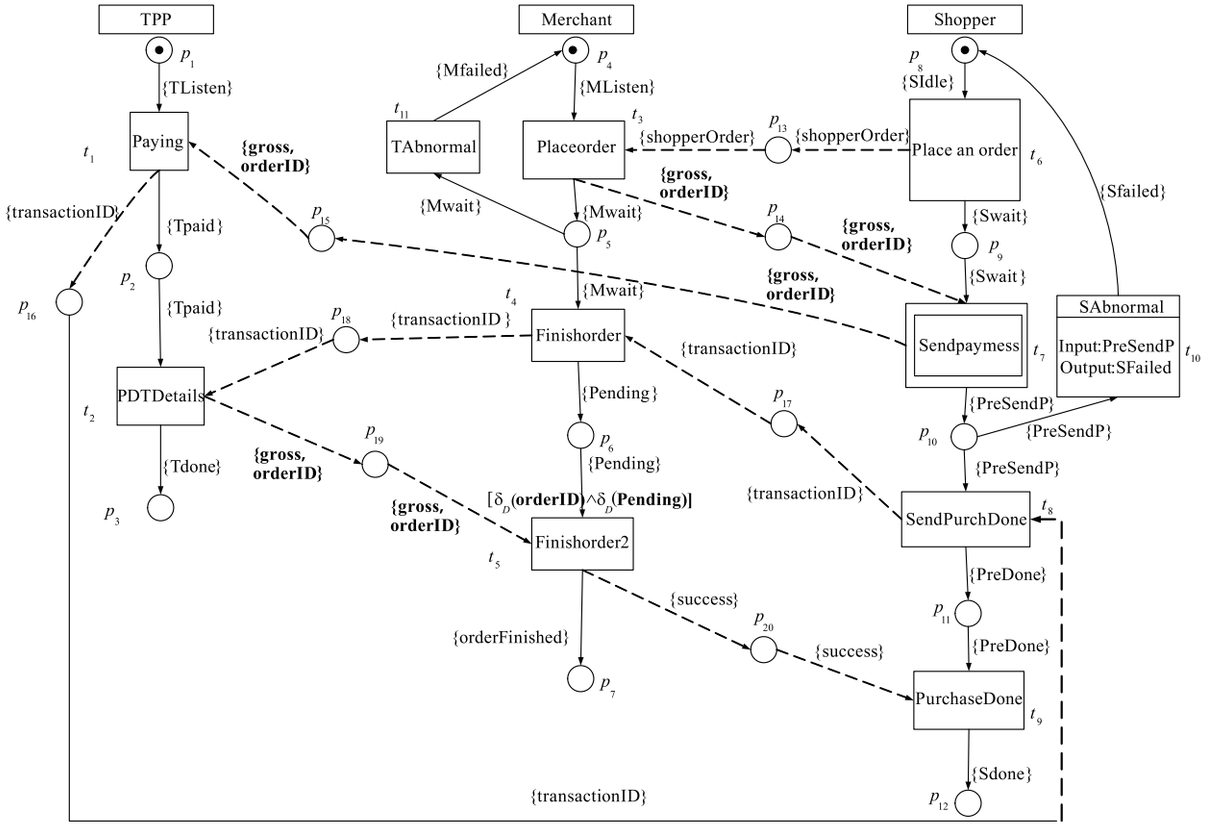

图3　一个较为完整的EBPN模型

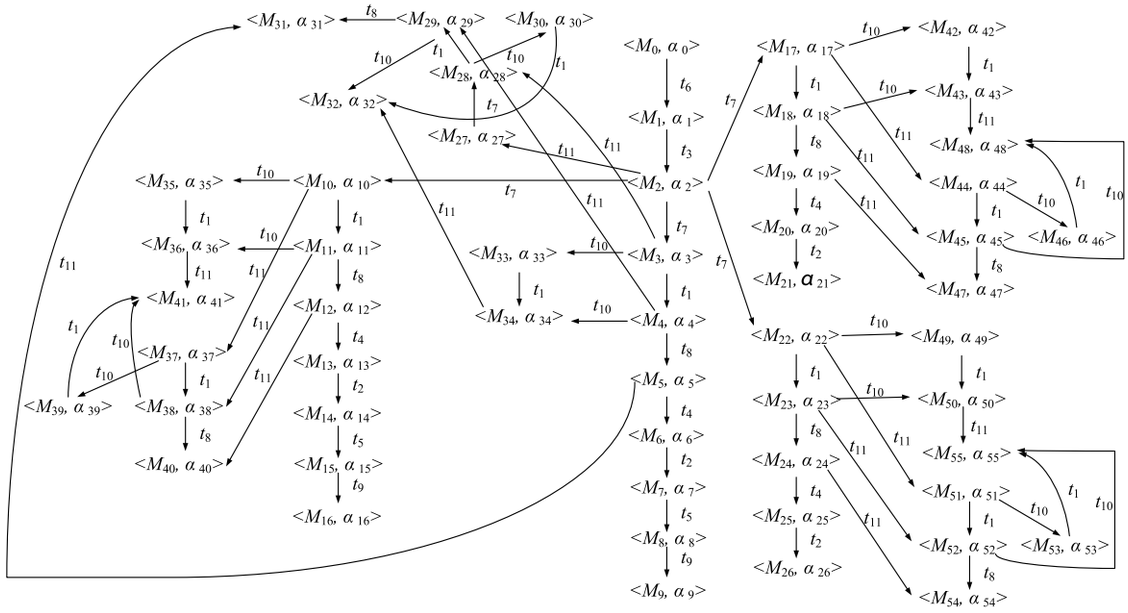

图4 图3所示EBPN模型的可达数据状态图

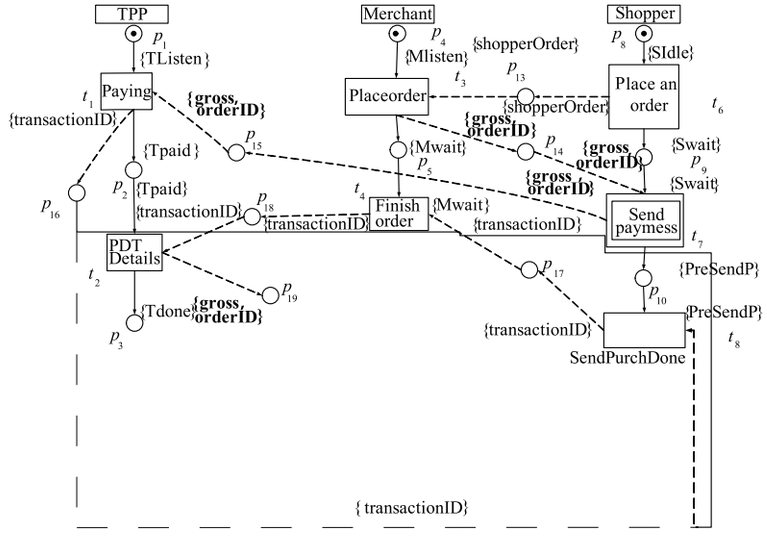

图5 图2所示EBPN模型的切片

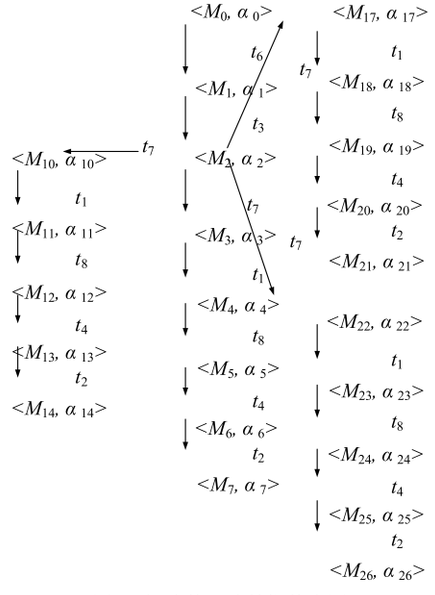

图6 切片的可达数据状态图

5 结束语

针对有界 EBPN 模型的可达问题,研究了可达数据状态图的构造方法以及相关分析结论. 借鉴程序切片的思想,研究了 EBPN 的切片方法,从而可用于降低 EBPN 可达分析的难度. 然而,对于某些 EBPN 模型,在极端情况下,状态爆炸问题仍然存在;因此,进一步研究 EBPN 模型的关联矩阵分析方法,以及 EBPN 模型某些子类的分析方法,是下一步需要研究的方向. 同时,开发基于 EBPN 的实用建模与分析工具也是下一步的工作重点.

参考文献

[1] 360 互联网安全中心. 2014 年中国网站安全报告[OL]. http://zt. 360. cn/1101061855. php? dtid = 1101062360&did = 1101205488,2015 - 03 - 20.

[2] 中国互联网络信息中心(CNNIC). 第 35 次中国互联网络发展状况统计报告[OL]. http://www. cnnic. net. cn/hlwfzyj/hlwxzbg/hlwtjbg/201502/t20150203_51634. htm,2015 - 02 - 03.

[3] 周展飞,周典萃,王贵林,卿斯汉. 电子商务协议的公平

性[J]. 电子学报,2000,28(9):13 – 15.
ZHOU Zhan-fei, ZHOU Dian-cui, WANG Gui-lin, QING Si-han. Fairness in electronic commerce protocols[J]. Acta Electronica Sinica,2000,28(9):13 – 15. (in Chinese)
[4] LEUNGWATTANAKIT W, ARTHO C, HAGIYA M, et al. Modular software model checking for distributed systems[J]. IEEE Transactions on Software Engineering, 2014,40(5):483 – 501.
[5] THOMAS K. State of Application Security Report[R/OL]. https://www. securityinnovation. com/comp-any/news-and-events/press-releases/state-of-application-security-report. html,2015 – 10 – 08.
[6] HOGLUND G,MCGRAW G. Exploiting Software:How to break code[M]. USA:Addison-Wesley Professional,2004.
[7] 黄颖,何克清,冯在文,黄贻望. 基于本体的业务流程适应性配置方法研究[J]. 电子学报,2016,44 (3):699 – 708.
HUANG Ying, HE Ke-qing, FENG Zai-wen, HUANG Yi-wang. Research on adaptive approach for business process configuration based on ontology[J]. Acta Electronica Sinica,2016,44 (3):699 – 708. (in Chinese)
[8] VAN-DER-AALST W M P. Service mining:using process mining to discover, check, and improve service behavior[J]. IEEE Transactions on Services Computing,2013,6(4):525 – 535.
[9] 毕敬,朱志良,范玉顺. Web 服务组合中行为兼容性分析与优化控制策略[J]. 电子学报,2011,39(12):2842 – 2849.
BI Jing, ZHU Zhi-liang, FAN Yu-shun. Behavioral compatibility analysis and optimal control policy in web services composition[J]. Acta Electronica Sinica, 2011, 39 (12): 2842 – 2849. (in Chinese)
[10] LIU Guan-jun. Some complexity results for the soundness problem of workflow nets[J]. IEEE Transactions on Services Computing,2014,7(2):322 – 328.
[11] DU Yu-yue, JIANG Chang-jun, ZHOU Meng-chu. A Petri net-based model for verification of obligations and accountability in cooperative systems[J]. IEEE Transactions on Systems, Man and Cybernetics Part A-Systems and Humans,2009,39(2):299 – 308.
[12] YU Wang-yang, YAN Chun-gang, DING Zhi-jun, et al. Modeling and validating E-commerce business process based on petri nets[J]. IEEE Transactions on Systems, Man and Cybernetics:Systems,2014,44(3):327 – 341.
[13] YU Wang-yang, YAN Chun-gang, DING Zhi-jun, et al. Modeling and verification of online shopping business processes by considering malicious behavior patterns[J/OL]. IEEE Transactions on Automation Engineering, 2014, DOI:10. 1109/TASE. 2014. 2362819.
[14] WANG Rui, CHEN Shuo, WANG Xiao-feng, et al. How to shop for free online-security analysis of cashier-as-a-service based web stores[A]. Proceedings of the 32th IEEE Symposium on Security and Privacy (SP)[C]. CA, USA:IEEE Press,2011. 465 – 480.
[15] LIU Guan-jun, JIANG Chang-jun. Secure bisimulation for interactive systems[A]. Proceedings of International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP)[C]. Switzerland:Springer,2015. 625 – 639.
[16] NEUMANN P. Principled Assuredly Trustworthy Composable Architectures[R]. SRI International Computer Science Laboratory,2004.
[17] 吴哲辉. Petri 网导论[M]. 北京:机械工业出版社,2006.
[18] HU He-suan, ZHOU Meng-chu, LI Zhi-wu. Liveness enforcing supervision of video streaming systems using nonsequential Petri nets[J]. IEEE Transactions on Multimedia,2009,11(8):1457 – 1465.
[19] LLORENS M, OLIVER J, SILVA J, et al. Dynamic slicing techniques for Petri Nets[J]. Electronic Notes in Theoretical Computer Science,2008,223(26):153 – 165.
[20] YU Wang-yang, DING Zhi-jun, FANG Xian-wen. Dynamic slicing of petri nets based on structural dependency graph and its application in system analysis[J]. Asian Journal of Control,2013,17(4):1403 – 1414.

作者简介

于汪洋　男,1983 年生. 博士毕业于同济大学电信学院,现为陕西师范大学计算机科学学院讲师. 主要研究方向为 Petri 网理论及应用、电子商务、可信软件.
E-mail:ywy191@ snnu. edu. cn

黄　昭(通讯作者)　男,1977 年生,英国布鲁内尔大学博士,加拿大渥太华大学博士后研究员. 现为陕西师范大学计算机科学学院讲师. 主要研究方向为社交网络、电子商务.
E-mail:zhaohuang@ snnu. edu. cn

方贤文　男,1975 年生,博士毕业于同济大学电信学院,现为安徽理工大学理学院教授,硕士生导师,主要研究方向 Petri 网、可信软件和服务计算.