安全关键信息物理系统的时序可预测性

李 曦[1] 孙贝磊[2] 万 波[2] 陈香兰[1] 周学海[1]
(中国科学技术大学计算机科学学院 合肥 230000)[1] (中国科学技术大学苏州研究院 苏州 215000)[2]

摘 要 安全关键的信息物理系统分为信息系统和被控的物理系统两个并发的子系统。其中信息系统具有分布式硬实时系统的特征,需要满足及时性和安全性要求,而时序可预测性是支持系统满足上述性质的关键。学术界对信息物理系统的时序可预测性的定义尚无统一认识。综述了信息物理系统的时序可预测性研究现状,总结提出了衡量系统可预测性的两个关键属性,包括时间可预测性和顺序可预测性。最后提出具有可预测性的信息物理系统的若干实现策略。

关键词 信息物理系统,时序可预测性
中图法分类号 TP301 **文献标识码** A

Temporal Predictability in Safety Critical Cyber Physical System

LI Xi[1] SUN Bei-lei[2] WAN Bo[2] CHEN Xiang-lan[1] ZHOU Xue-hai[1]
(School of Mathematics and Computer Science, University of Science and Technology of China, Hefei 230000, China)[1]
(Suzhou Advanced Institute, University of Science and Technology of China, Suzhou 215000, China)[2]

Abstract The safety critical cyber physical systems (CPS) consists of two concurrent subsystem, including cyber systems (CS) and physical systems (PS). The CS is constructed as a distributed real-time systems to meet the requirements of timelines and safety, both of which can be supported by the key property in system, i. e. , temporal predictability. However, there are no clear definition of temporal predictability in CPS. This paper summarized the state-of-art of predictability in CPS. Two key attributes to improve the system's predictability were summarized, including timing predictability and ordering predictability. Finally, some advice to build a predictable CPS were proposed.

Keywords Cyber physical cystem, Temporal predictability

1 引言

典型的安全关键信息物理系统(Cyber Physical Systems, CPS)包括汽车、飞机、空中交通控制、电网、炼油、医疗设备、患者检测以及智能建筑等,其系统复杂性表现为具有规模大、不确定性和动态性等特性。CPS系统整体上可分为信息系统(Cyber Systems, CS)和被控的物理系统(Physical Systems, PS)两个并发的子系统。其中CS系统是计算、通信和控制等子系统的有机集成和协同,是一个典型的分布式实时系统。一方面,CS对PS的感知和操纵发生在本地节点,其整体行为呈现为节点间的消息通信或其他形式的交互过程;另一方面,CS的正确性不仅取决于计算结果的正确性,更取决于产生结果的时间和顺序的正确性。CS的实现不在于追求操作的平均性能,而在于功能和时间行为的可预测性(Predictability)保证。

CPS系统具有反应式系统特征。CS不断地对PS的状态进行采样输入,计算状态变化和输出相应的控制。这一过程需要与PS的实时状态变化保持同步,以保证CS的响应与PS的当前状态之间的因果关系(输入输出事件顺序)。响应时间是反应式系统的核心指标之一,一般依赖于系统的当前负载情况。为了保证系统的响应时间,需要对系统的环境输入负载进行预测,如输入事件的到达时间间隔是周期性的还是非周期性的,到达的事件是否存在最小到达间隔,是否存在事件突发到达,是否存在突发到达最小间隔,一次突发到达的最大事件个数等。

系统行为的可预测性非常重要。控制理论的基础原理指出,如果给定系统的初态、输入和时间间隔,基于一个确定性的物理模型,可以计算出控制系统的当前状态,或者说其系统行为是可预测的。然而,对于CS的一段特定程序代码,如果给定程序的初态、输入和期望的执行时间,目前主流的编程范式无法可靠地预测程序的未来状态。这使得CPS系统设计非常困难。这种系统只能针对特定应用的设计约束,依赖于特定硬件架构和软件环境,且需要精确的调试,非常脆弱,可靠性差。微小的操作条件或硬件平台的变化都可能造成系统行为的巨大改变。

不确定性一方面源于CPS系统底层的感知层的内在特性,如采样的时空限制(物理系统中事件发生的时间间隔小于信息系统可观测的间隔),另一方面源于系统时间属性的不确定性(Uncertainty),如时钟准确性有限和网络延时抖动[2]。由运行时系统的任务调度机制和共享资源管理策略所决定的任务执行顺序变化等也是不确定性的重要来源。Liu指出"不可抢占调度和竞争总是引入执行的不可预测性"[3]。共享资源竞争可能引发任务执行的优先级翻转,导致违反时间约束。另一个重要因素是系统体系结构造成的程序执行时间抖动,如流水线乱序执行、Cache失配或置换以及系统总线竞争都可能导致程序执行时间不确定。

可预测性保证了对物理系统响应的及时性,简化了系统事件因果分析的复杂度,有利于提升系统的可测性(Tes-

本文受国家自然科学基金项目(61272131)资助。

李 曦(1963—),男,博士,副教授,主要研究方向为嵌入式系统、可重构,E-mail:llxx@ustc.edu.cn。

tability)。主动冗余容错技术要求各复制节点的行为是可预知的。可预测性也保证了基于构件的设计方法中所必需的可组合性。

STANKOVIC 很早就提出了“可预测实时计算”的概念[1]。但学术界对信息系统可预测性的确切含义和形式化定义一直未达到共识，可预测性理论和实现技术非常不完备。可预测概念一方面被用于描述设计方法学的特征，另一方面作为系统的一个定性属性。这种状况使得现有的硬实时系统设计方案具有极强的专用性，往往只能被动地选择具有充分的(而不是足够的)时间可预测性的组件，大大增加了系统的实现成本。

因此，研究时序可预测问题对保证安全关键 CPS 系统的行为正确性和可靠性十分必要。我们认同“可预测性是系统的内在属性，是系统不确定性的度量”的观点，并且认为安全关键的 CPS 系统的可预测性不仅包括时间可预测，还应包括顺序可预测性。本文提出时序(Temporal)可预测性概念，包括定时(Timing)行为可预测性和定序(Ordering)行为可预测性两大类，涉及进程的执行时间界和网络通信延时界的稳定性(Steadiness)和紧致性(Tightness)，以及任务执行与事件处理顺序的乱序和一致性程度等特征。在此基础上，进一步给出 CPS 系统可预测性的分级判据。通过形式化定义，可预测性成为可度量和可比较的系统指标。最终根据这些指标给出构建具有可预测性信息物理系统的若干策略。

2　相关研究

实时系统必须保证有界性和响应时间的可预测性[6]。概念上如果可以预言一个系统在未来什么时间将发生什么事情，则此系统行为是可预测的(Predictable)。Ortega 指出“语义完整性和及时性(Timeliness)构成可预测性”[5]；Douglass 定义可预测性为“系统的响应特性所达到的可预知程度”[8]；Henzinger 认为可预测性是确定性的一种类型[9]；Jensen 认为可预测性是一个连续值，其最大值为确定性(Determinism)，其最小值意味着没有任何东西可以预知[10]。可预测性(或不确定性)的度量以及最小可预测性(最大不确定性)的特征依赖于特定的可预测性模型。确定性意味着一致性(Consistent)和可重复性(Repeatable)。Lee 认为可预测性是预期一个系统的行为的能力，是布尔属性[11]。因此，可预测性是需要证明的，即为了预期某种属性，需要某个终止过程以显示保持了该属性。Koptz 认为，如果系统随时间变化的行为是可预测的，则此系统是确定的(Deterministic)[4]。

研究者尝试对可预测性和时间预测性的准确含义和形式化定义进行讨论。Liu 从一个作业集合的执行行为角度分析了可预测性，指出对于给定的优先级驱动调度算法，如果作业集 J 中的每个作业在实际调度中的实际开始和完成时间都限定在最大和最小调度的开始和完成时间内，则 J 的执行是可预测的，也即开始和完成时间可预测[3]。

Reinhard 小组为了改进程序的最坏执行时间(WCET)分析方法，对时间(Timing 或 Time)可预测性的定义和度量进行了集中研究，并于 2004 年首次尝试量化定义时间可预测性[13]。他们以真实 WCET 与 WCET 界的时间间隔(所谓悲观分析，见图 1)度量一个程序的时间可预测性，其值越小，则可预测性越好，当最好执行时间 BCET 等于 WCET 时，认为可预测性最高。此定义中也考虑了一个程序的执行时间下界和 BCET 的差值。Grund 进一步给出了一个可预测性定义的模板，包括 3 个方面：待预测的属性、不确定性来源、量化指标[14]。在此基础上，他们将时间可预测性定义为最小执行时间除以最大执行时间的商，商为 1 的系统可预测性最好，并认为每个系统都有一个内在的可预测性因子，WCET 分析不应是这个因子的一部分。与其他研究的不同在于，Reinhard 等认为可预测性是系统的内在属性，希望以此为指标对不同系统进行比较。然而，真实的 BCET 和 WCET 通常是未知的，因此，Martin 认为使用该定义无法测量时间可预测性，对系统设计与比较[15]无益。他们认为对于一个实际系统而言，仅仅考虑时间可预测性属性是不完备的，还需要考虑性能，也就是最坏情况下的性能(WCET 界)；他们认为比较两个系统的时间可预测性时应该比较它们的 WCET 界。而一个所谓的“系统”应由硬件、编译器和 WCET 分析工具构成，确定 WCET 的界依赖于这三者。

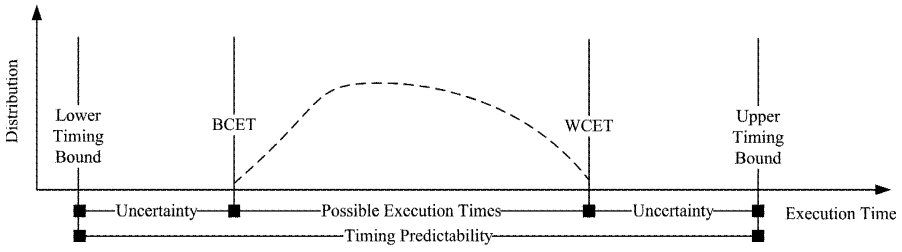

图 1　程序执行时间分布

Puschner 等提出“时间可预测计算”概念，定义一个系统模型的时间可预测性为计算实际系统中活动的持续时间的能力[7]。所谓“能力”指难易程度，而不是可判定性。因此，为了提高时间可预测性，不仅需要系统时间行为计算的准确性，也需要高效的计算方法。定义可预测性为可分析性和稳定性的乘积(包括权重)[7]。稳定性是时序模型的 BCET/WCET 的商。对 WCET 可预测性而言，稳定性被忽略，其因子设为 0，此时，可分析性等于 WCET 可预测性，即 BCET 除以 WCET 界的商。

综上，多数研究者认为可预测性等价于确定性，论述时往往将两者混用，如 Guenter 在标题中使用“时间可预测”，而在结语中却使用“时间确定”[16]。其次，上述研究都是从单个计算任务的执行时间分析角度定义时间可预测性，而 CPS 系统的时间行为包括输入输出、计算与通信。另外，现有研究都关注定时属性，而忽略了定序属性。任务或事件执行顺序一方面是系统的功能约束，另一方面也影响系统时间行为的正确性。可预测性不仅应包括时间可预测，还应包括顺序可预测性。

3　可预测性的定义与度量

时间是安全关键 CPS 系统的关键属性，为预测、验证和控制 CS 系统行为提供了重要的参考信息。PS 系统采用连续

时间模型刻画其行为,CS 系统则采用离散时间模型(或称致密时间模型)。CS 系统通常采用带有时间戳的事件刻画描述系统行为,根据时间戳信息对事件排序,确定任务执行顺序或事件处理顺序,以此提高事件对应的执行过程的时序可预测性。

定时行为涉及时间上下界,定时可预测性指时间上下界的稳定性和紧致性。定序行为涉及序列顺序的乱序程度,定序可预测性可由观察一次执行所得到的事件序列的乱序程度来刻画。

3.1 时间可预测性

任务是响应事件的实体。任务的执行约束包括定时约束、优先约束和资源约束等,其中定时约束包括到达时间、释放时间、开始时间、完成时间、响应时间以及截止时间。如果单个处理器上需要执行多个并发的任务,则需要实时调度算法决定何时且按何种顺序调度这些任务,以保证它们在各自的截止时间之前完成。实时调度理论往往假设各个任务的执行时间已知,但实际它不是一个常数(见图 1),取决于输入、系统状态和硬件的定时行为,安全起见只能使用 WCET。准确的 BCET 和 WCET 很难获得,只能估算执行时间的上下界。估算一般采用测量或静态分析两种方法,估算结果好坏的标准是安全性(预估的 WCET 要大于或等于实际的 WCET)和 BCET、WCET 与上下界之间的紧致性。任务的执行时间可预测性的具体定义可参考文献[13],任务的响应时间可预测性定义可参考文献[17]。

消息通信是分布式系统各节点间协同的基本机制。消息传递(deliver)时间指消息 m 的发送事件时刻与节点 p 中对应进程收到 m 事件时刻之间的间隔。由于网络传输的延迟和各节点进程执行速度的异步性,导致消息传递时间不确定。如果消息传递时间有界,则称此通信协议为同步通信协议,反之则为异步协议,相应的分布式系统被称为同步系统或异步系统。因此,消息传递时间可预测性可由传递延时的稳定性和紧致性定义。稳定性定义为同一节点消息传递时间的最大差值,紧致性定义为向任意两个节点发送消息 m 的最大传递时间之差。本文将传递时间可预测定义为稳定性与紧致性的乘积。

3.2 顺序可预测性

任务执行顺序发生变化可能导致系统功能失效。首先,对于单处理器系统,假设有 5 个独立的进程,它们按非抢占的方式执行,则有 120 种不同的执行顺序(如果在多处理器系统上或在可抢占的系统上,则有无穷多种交叉执行组合)。在一个常规的并发系统中(假定任务独立,没有优先约束),不一定要规定进程执行的精确顺序。如果程序是正确的,那么,无论其内部行为或实现细节如何,程序的功能性输出将是一样的。然而,即使程序的输出结果在所有可能的交叉执行情况下都相同,输出结果的时刻却无法保证一致。如果上述 5 个进程中某个进程的时限很紧迫,那么也许只有这个一进程首先执行才能满足系统的时间性需求。其次,对于多处理器或多核系统,任务执行时间或顺序的变化有可能造成调度异常问题[12],即调度时长异常增加。此外,对于抢占式调度系统,虽然设计时已按优先级预定任务的执行顺序,但运行时仍可能因资源竞争造成优先级翻转,从而出现设计时和运行时任务执行顺序不一致的问题。由上,时序不正确有可能导致安全关键的 CPS 系统的功能失效。因此在系统运行时监测中,亟需分析任务执行顺序,及时发现系统中的极端异常行为。

事件定序是分布式系统研究中的重要问题。不同于状态模型(或称动作模型)关注组件内部状态及其变迁,事件模型强调组件与组件之间或组件与被观察对象之间的交互、协同和组合。事件是一个重要的抽象概念,系统赋予了事件多重语义,如:

(1)反应式系统的事件动作模型中,事件标识了系统在时间或空间上发生的一次状态变化。事件触发动作,动作改变状态[18]。

(2)分布式系统的协同模型中,事件为进程间或角色间(actor)的消息收发动作[19]。

(3)并发系统的进程演算模型中,事件为两个进程间共享的一个动作[20]。

按事件自身的性质划分,事件包括原子事件、复杂事件、瞬间事件、持续事件等类型。按事件的功能划分,CPS 事件包括外部事件(输入输出)、内部事件(组件通信消息)。外部事件为 PS 与 CS 系统间的交互行为。内部事件发生于某个进程内部,对其他参与者不可见。系统可通过发送事件、接收事件以及事件传输描述各软件组件间的通信和协作关系。

事件定序指按照某种特定的规则建立事件流中各个事件的顺序关系。顺序关系包括两大类:全序和偏序。定序算法包括无序算法、FIFO 算法、因果算法和全序算法等。这些算法的主要区别在于所允许的异步执行的程度不同。另外,事件模型所基于的时间模型包括因果序模型、离散时间模型和连续时间模型等。

事件流的顺序关系描述了 CPS 系统的行为模式。事件触发是 PS 与 CS 交互的基本模式,CS 对 PS 的响应由分散于各个节点的功能模块(称为“进程”)协同实现。由于 PS 与 CS 的时间模型不同,PS 中发生时刻非常接近的事件,在 CS 中可能无法恢复其真正的时间顺序,甚至与 PS 中的时间顺序矛盾。为了保证整个 CS 系统行为的一致性,要求各个节点按相同的顺序对外部事件进行响应,并与这些事件发生的时间和顺序保持一致。

对于分布式容错技术而言,丧失相互备份通道(replicated channel)的确定性意味着通道的错误屏蔽能力失效[4]。为了实现冗余容错系统行为的备份确定性(replica-deterministic),各节点间的消息传递系统必须是可预测的,即消息传递的时刻需要预知,且接收消息的顺序需要与所有通道发送消息的顺序保持一致。

基于上述认知,Sun 等基于逆序数理论首次定义了任务执行顺序和输入就绪顺序的可预测性[17]。这些定义明确了 CPS 顺序可预测性,且其提供的量化方法可作为分析和比较系统的顺序可预测性的参考指标。

3.3 时序可预测性分级

依据上述对定时和定序可预测性的定义,可进一步将可预测性划分为如下 3 个级别。

(1)可预测:定时指标有界,且界已知;按序,且序一致。

(2)部分可预测:定时指标有界,但界未知,或某些情况下有界;按序,且序一致。

(3)不可预测:定时指标无界,且乱序。

4 可预测性实现策略

实现可预测 CPS 要求进程的每一步执行时间有上、下

界，网络通信延时有界和本地时钟偏差率有界。现代处理器所采用的 Cache、TLB、流水线和分支预测等技术，都使得程序的执行时间具有上下文依赖性。单个指令的执行时间甚至一个存储器的访问操作都依赖于执行的历史状态。现代操作系统中的调度技术、线程技术等是基于“尽力而为”的思想，仅适用于静态的、不存在或仅存在少量并发行为的实时系统。CPS 是动态的和高度并发的。随着 CPS 复杂性的增长以及多核处理器技术等的引入，操作系统进程之间、处理器核之间的通信越来越频繁，进一步降低了系统时序行为的可预测性。Simulink，Modelyze 和 PtolemyII 等许多系统建模平台都包含精确建模和控制时间的手段，但几乎都不能提供任何时序正确性保证。设计和实现具有时序可预测性的 CPS 系统仍然是一个巨大的挑战，本文建议考虑如下原则。

(1)采用支持定时和定序语义的编程语言。一类是具有 at，within，until，every，after 等语言结构的异步语言，通过定义与平台无关的定时行为使得程序具有时间确定性；另一类是基于逻辑时钟的同步语言。逻辑嘀嗒用于事件定序，无须与实际物理时间相关联。

(2)采用时间触发的系统执行模式。CS 对外部事件的响应有两种基本方式：事件触发(ET)和时间触发(TT)。ET 方式中一旦输入事件发生(到达)就立即触发系统响应。TT 方式中响应外部事件的时刻由系统时钟控制，与输入事件是否到达无关；系统输出也与时钟同步，无论输出产生多快。ET 中输出抖动源于处理时间的变化，而 TT 中则由时钟精度决定。ET 灵活且资源利用率高，但 TT 的时序可预测性优于 ET，可在编译时通过静态分析确定定时约束是否满足。TT 假设 WCET 已知，需要提前指定时态控制结构，包括开始时间和完成时间。实际应用中两种模式可组合使用，如采用 ET 进行输入，按 TT 进行输出。

(3)采用消息定序协议(message ordering protocol)和时间触发通信协议。

(4)采用具有时序可预测性的硬件平台。

(5)采用稀疏时间结构(sparse time base)作为系统的时间基，以固定顺序执行一段时间内的并发任务，或基于事件排序确定任务的执行顺序。

此外，CPS 系统并非总需要具有很高的时间可预测性。对于分布式协同而言，只要能保证具有因果关系的事件的执行顺序，就保证了系统行为的正确性。对于没有因果关系的并发事件，可以按任意顺序执行。

结束语　可预测是实现可控、可重复、可调试和可仿真的前提。可预测的系统可信度高，实现成本低。我们认为对具有实时、反应式和分布式特征的安全关键 CPS 系统而言，由定时和定序特征构成的“时序(Temporal)可预测性”概念更加完备，它包括事件的响应时间、任务的执行时间以及事件的响应顺序和任务的执行顺序等性质。基于此概念构建具有时序语义一致性的编程模型、分布式实时操作系统和硬件架构是实现可预测 CPS 的重要基础。

参 考 文 献

[1] STANKOVIC J，RAMAMRITHAM K. What is predictability for real-time systems? [J]. Real Time Systems，1990，2(4)：247-254.

[2] KOPETZ H. Temporal uncertainties in cyber physical systems [M]. Advances in Real-Time Systems. Springer Berlin Heidelberg，2012：27-40.

[3] LIU J W S. Real-time systems [M]. Prentice Hall，2000.

[4] KOPETZ H. Real-time systems：design principles for distributed embedded applications [M]. Springer Science & Business Media，2011.

[5] ORTEGA R. Timing predictability in real-time systems [J]. Univ. of Washington，1994.

[6] BUTTAZZO G. Hard real-time computing systems：predictable scheduling algorithms and applications [M]. Springer Science & Business Media，2011.

[7] KIRNER R，PUSCHNER P. Time-predictable computing [C]// IFIP International Workshop on Software Technolgies for Embedded and Ubiquitous Systems. Springer Berlin Heidelberg，2010：23-34.

[8] DOUGLASS B P. Doing hard time：developing real-time systems with UML，objects，frameworks，and patterns [M]. Pearson Schweiz Ag，1999.

[9] HENZINGER T A. Two challenges in embedded systems design：predictability and robustness [J]. Philosophical Transactions of the Royal Society of London A：Mathematical，Physical and Engineering Sciences，2008，366(1881)：3727-3736.

[10] JENSEN E D. Wrong assumptions and neglected areas in real-time systems[C]// 2008 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC). 2008.

[11] LEE E A. Predictability，Repeatability，and Models for Cyber-Physical Systems [C]// Workshop on Foundations of Component Based Design(WFCD) at ESWeek 2010. 2010.

[12] GRAHAM R L. Bounds on the performance of scheduling algorithms [J]. Computer and Job Scheduling Theory，1976：165-227.

[13] THIELE L，WILHELM R. Design for timing predictability [J]. Real-Time Systems，2004，28(2/3)：157-177.

[14] GRUND D，REINEKE J，WILHELM R. A template for predictability definitions with supporting evidence [C] // OASIcs-OpenAccess Series in Informatics. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik，2011：18.

[15] SCHOEBERL M. Is time predictability quantifiable? [C]// 2012 International Conference on Embedded Computer Systems (SAMOS). IEEE，2012：333-338.

[16] KHYO G，PUSCHNER P，DELVAI M. An operating system for a time-predictable computing node [C] // IFIP International Workshop on Software Technologies for Embedded and Ubiquitous Systems. Springer Berlin Heidelberg，2008：150-161.

[17] SUN B，LI X，WAN B，et al. Definitions of predictability for Cyber Physical Systems [J]. Journal of Systems Architecture，2016，63：48-60.

[18] MARINESCU D C，LUMPP J E，CASAVANT T L，et al. An Event-Action Model and Associated Architecture for Monitoring Parallel and Distributed Systems[J]. Computer Science，1988.

[19] ACTORS A G. A Model of Concurrent Computation in Distributed Systems，Series in Artificial Intelligence[M]. MIT Press，1987.

[20] BROOKES S D，HOARE C A R，ROSCOE A W. A theory of communicating sequential processes [J]. Journal of the ACM (JACM)，1984，31(3)：560-599.