面向汽车信息物理系统的 STPA 方法研究的文献综述

曾珂瑜 S181000893

（湖南大学 信息科学与工程学院，湖南 长沙）

摘要：随着智能网联汽车的出现，系统的物理过程与网络相互交融。除了保障汽车功能安全外，汽车的信息安全近些年也得到广泛关注。为尽量避免由功能安全或者信息安全被影响而导致的损失，必须进行危险分析。传统的分析方法大都只注重于评估功能安全，已不再适用于复杂的汽车信息物理系统，因此引发了系统理论过程分析方法的出现，基于这一方法，逐渐衍生出了多种新的危险分析方法，用于同时保证功能安全和信息安全，完善了基础的系统理论过程分析方法。

关键词：汽车信息物理系统、危险分析方法、功能安全、信息安全

0 引言

汽车变得越来越智能化，我们将进入智能网联汽车时代，必将引发一场深刻的技术革命。信息物理系统（CPS）将物理过程和通信技术紧密联系在一起。随着计算力和网络传输速度的增加，工业控制系统中产生了越来越多的、以信息通信技术为基础的应用，更加强了物理过程和网络领域的联系。因此也使得网络更加互联，处理更加复杂。危险分析技术必须同时用于保障功能安全和信息安全。

传统的危险分析方法多注重于评估功能安全，已不再适用于智能网联汽车的复杂电子系统。系统理论过程分析 STPA 作为一种新型的危险分析方法，最初应用于航天领域中，它能通过发现系统控制结构中存在的设计缺陷，找出不恰当的控制行为，进而避免系统事故的发生。此外还能分析系统中存在的设计问题，找出系统危险的根本原因，改进系统设计。STPA 能用来进行汽车功能安全危险分析，在 STPA 基础上发展的一种改进方法能够实时地进行危险分析，另一种改进方法 STPA-Sec 能够用于同时评估功能安全与信息安全，有望解决智能网联汽车缺乏功能安全和信息安全综合的一体化保护体系的问题，同时有望实现针对驾驶自动化、网络攻击、AI 算法输出不确定性等异常行为的统一自适应处理。

1 危险分析技术产生背景

汽车的电子化、智能化和网联化发展使得它所包含的电子系统复杂性骤增。一些高端车型包含的代码量极大。同时由安全气囊、电子稳定性控制和紧急制动辅助等安全相关电子功能失效引发的功能安全问题，导致了一系列安全事故，并受到了社会各界的持续关注。因此，功能安全已成为智能网联汽车进入实用发展的核心指标之一，亟需遵循 ISO 26262 和 GBT 34590 标准，从设计之初就将功能安全保障作为下一代智能网联汽车设计的首要要求。

2006 年信息物理系统（Cyber-Physical Systems, CPS）这一概念的催生[1]将物理过程和信息与交流通信技术紧密联系在一起。如何确保功能安全和信息安全成为目前 CPS 发展的重要关注点[2]。下一代智能网联汽车将面临着全新的信息安全挑战。例如智能驾驶决策错误、图像识别算法不可靠、AI 算法输出的不确定性以及信息安全入侵攻击等。随着汽车终端信息技术的发展以及无线接口的增多，智能网联汽车的安全漏洞也随之增多。近年来，汽车信息安全问题愈发严峻，由汽车信息安全导致的召回事件更是引发社会各界的高度关注。下一代智能网联汽车实现了驾驶辅助功能，但仍发生了多起安全事故，造成了人员伤亡和财产损失。智能网联汽车的信息安全漏洞也随之激增。攻击者可能通过漏洞入侵汽车内部，以信息篡改、病毒入侵和恶意代码植入等手段对汽车进行网络攻击，恶意访问汽车内部敏感信息，甚至恶意操控汽车，威胁人身安全[3]。这些由信息安全破坏引起的损失和危险逐渐受到广泛关注。根据 SAE J3016《标准道路机动车驾驶自动化系统分类与定义》中自动驾驶分级情况可以看出，智能网联汽车的出现使得驾驶主体从人转变为人机交互[4]，信息安全攻击不仅将造成隐私泄露或财物丢失，而且极有可能危及驾乘人员和周围行人的生命安全，甚至上升为国家公共安全问题。因此必须采取有效措施来实现下一代智能网联汽车的信息安全防护和保障，下一代智能网联汽车的信息安全保障势在必行。

现有行业的关键基础设施是软件、网络、机电构件和人组成的复杂工业控制系统。传统的危险分析技术假设事故是由组件失效引起的[5]，并且过分简化了人的作用[6,7]。此外，大多数与软件相关的事故都可以追溯到软件约束信息不完整或存在缺陷这些原因上来[8,9]。一些专家认为：安全事故或灾难是复杂系统状态的“涌现”；基于部件失效或数值超限概率模型和归因理论，无法全面揭示事故

缘由并提出对策；传统的危险分析方法无法应对“软件密集和人参与控制”型复杂系统的安全性分析要求。因此，可以将事故的起因看作是在系统设计、开发、运行中不恰当控制或安全相关约束不充分的执行导致的，看作是由组件间交互而导致违背系统安全约束的结果。从计算机的视角来看，组件间交互需要处理器，且需要通信等资源，这样便会出现资源竞争，而造成没有竞争到资源的组件无法正常的执行，进而造成安全约束无法满足。此外，由于组件之间的互联性，信息安全已经不能简单同功能安全混为一谈，它不仅可能因隐私信息被盗等欺诈行为造成财产损失，还可能产生人员伤亡，严重时可能威胁国家的利益，引起社会的恐慌。因此，对汽车 CPS 的功能安全和信息安全进行正确评估、对危险采取合理的防护措施、采取“早分析早预防”的机制迫在眉睫。在定义组件障碍场景后，工程师们就可以使用容错或故障安全技术，以防止所识别的故障所造成的危害，并提高单个组件的完整性[10]。由此可见，为尽量避免由功能安全或者信息安全被影响而导致的损失，必须进行危险分析。

2 传统危险分析方法以及 STPA 的出现

常见的危险分析技术主要有两类，其一是基于故障的危险分析技术，另一类是基于系统的危险分析技术。其中，事故分析树 FTA[11]和失效模式与影响分析 FMEA[12]都是典型的基于故障的危险分析技术。事故分析树 FTA 首先确定事故结果，将要分析的事故作为顶上事件，层层追溯，上层事件是下层事件的必然结果，下层事件是主要原因，上下层间用逻辑门连接，直至找出发生事故的最基本原因为止。这样形成一棵以事故结果为根，以原因事件为枝干的倒立逻辑树。其主要遵循的程序有分析准备阶段、编制事故树阶段、定性分析阶段、以及定量分析阶段四部分。失效模式与影响分析 FMEA 其是在产品设计阶段和过程设计阶段，对构成产品的子系统、零件，对构成过程的各个工序逐一进行分析，找出所有潜在的失效模式，并分析其可能的后果，从而预先采取必要的措施，以提高产品的质量和可靠性。其分类之一设计 FMEA 在一个设计概念形成之时或开始之前进行分析，过程 FMEA 在生产准备之前、在过程可行性分析阶段或之前开始。文献[13]通过研究复杂系统危险和故障率的计算方法，并考虑时间和其他的特定条件下对比了 FTA 和 FMEDA 等方法。危险和可操作性分析 HAZOP 用于定义

系统是如何偏离设计行为的，是一种基于系统的危险分析技术[14]，在设计过程中，其通过对每一个细节提出问题，采用“关键词”高效地发现设定参数与实际的偏差，由此找出出现偏差的原因以及可能产生的影响。因为现代系统越来越依赖软件系统来控制物理过程，并进一步嵌入到社会技术环境中，所以该方法很难量化。相反，基于故障的危险分析技术的结果虽然可以量化，但其并不适用于组件互联导致的事故。Avizienis 等人于 2004 年提出了“fault-error-failure”链式模型[15]，对于复杂的 CPS 并不适用。

上述所有的传统危险分析方法都未能涵盖新出现的网络物理系统的复杂性。他们只能用于评估单个组件的线性系统，而智能网联汽车中的危险状态不仅可能是组件失效引起，还可能是组件交互导致的冲突引起的。针对于这一不足，Leveson 提出了一种系统理论事故模型与过程（STAMP）分析方法[16]。在系统理论事故模型与过程 STAMP 中，系统安全被视为控制问题。其基于系统理论分析，在开发、设计和运行阶段中存在地控制问题。当这些控制问题满足安全约束时，才能保障系统安全性，否则，会导致相应的系统事故发生。此外，系统理论事故模型与过程 STAMP 还给出了 4 类不恰当的控制类型：无控制行为、错误的控制行为、正确的控制行为开始的时间过早或过晚、正确的控制行为结束的时间过早或过晚。以系统理论事故模型与过程为基础，Leveson 接着提出了系统理论过程分析 STPA 和 CAST 事故分析。系统理论过程分析技术注重于怎样从设计中找出不准确的控制，而 CAST 事故分析技术重在研究如果找到事故发生的不准确的控制。还引出了基于 STamP-based 的分析[17]事故重建的方法[18]以及其他的一般安全分析的方法。STPA 最初应用与航空航天领域，飞机的软件是安全的关键，因为它能监视和控制可能涉及到危险系统行为的各类组件[19]。这对于汽车 CPS 来说也是一样的。系统理论过程分析 STPA 旨在通过一种新的基于系统的危险分析技术来应对基于故障的危险分析技术和基于系统的危险分析技术所面临的新的挑战。主要通过定义分析的目的、建立控制结构、识别不安全的控制行为、识别致因场景四个方面进行分析。虽然很难量化 STPA 的结果，但 Thomas 在[20]中提供了一个系统理论过程分析 STPA 的数学模型和一个系统地进行系统理论过程分析的过程，可以用来量化 STPA 的结果。

与传统的危险分析技术相比，系统理论过程分析方法 STPA 可以用于分析非

常复杂的系统；可以在早期概念分析中启动，以帮助确定安全要求和制约因素，并用于增加系统架构和设计的安全性，从而消除在后期或运行期间识别出设计缺陷时的昂贵返工。其能通过发现系统控制结构中存在的设计缺陷，找出不恰当的控制行为，进而避免系统事故的发生。还能分析系统中存在的设计问题，找出系统危险的根本原因，改进系统设计。此外，系统理论过程分析方法 STPA 在时间和资源方面的费用也要低得多。由于传统的系统理论过程分析方法主要用于评估功能安全，而网络攻击导致的信息安全隐患同时也可能影响功能安全。因而需要找到网络攻击和物理过程之间的关系，将功能安全与信息安全同时纳入考虑。相比而言，STPA 发现了上述方法所能发现的任何因果方案，同时还能发现非单组件故障的场景以及软件相关的场景。除了适用范围更广以外，STPA 无论是在资源的使用还是时间的花费上，都要明显比传统危险分析方法少。

总体而言，STPA 的优势如下：

（1）启动时间早。STPA 可以在系统概念分析的早期启用，可以在系统设计开始前完成故障的分析，确定系统应该满足的约束条件；

（2）适用范围广。STPA 适用于非常复杂的系统的危险分析，传统方法在系统设计或运行中才能发现的问题，STPA 可以在设计早期找出问题，并提出相应的缓解策略；

（3）考虑因素全面。使用 STPA 进行危险分析，除了传统方法能找到的损失因果因素外，还将系统软件和操作人员考虑在内，能发现更全面的损失因果因素；

（4）中间结果清晰。STPA 是一种有明确步骤的分析方法，经过每一步的分析，都会产生相应的中间结果，可以在大型复杂系统设计中作为参考文档。

基础的 STPA 流程简单，可操作性高，但也具有一些缺点。一方面，STPA 应用于系统设计确定、完善以及扩充之前，但穷举可能出现的故障是非常难的，一旦在系统的设计过程中，发现了先前没能发现的问题，整个分析工程都需要返工，影响分析效率。为解决这两个问题，两个基于 STPA 的新型方法被提出。其一是 John Thomas 等人提出了一种基于 STPA 的改进方法[21]，将开发任务和危险分析任务集成在一起。随着系统设计需求变化，分析人员可以很快发现新的危险，并将安全隐患立即反馈到工程流程中，从而减少额外的返工，提高了危险分

析的效率。

另一方面，基础的 STPA 只将功能安全考虑在内，而在智能网联汽车的危险分析工作中，信息安全和功能安全是同等重要的。基于这个问题，研究者们对于网络攻击对信息安全的影响进行了深入的探究。Dondossola 等人在文[22]中研究了网络攻击对信息和通信技术通信能力的潜在威胁。Wang 和 Lu[23]强调了对可用性、完整性和保密性的网络攻击以及它们对不同用例的潜在影响，并进一步提出了潜在的缓解策略。此外还提出了将签名和加密作为密码对策，并解释了有限的计算能力和严格的时间限制所带来的困难，该分析很好地概述了 CPS 所需要的安全考虑。Kundur 等人在文献[24]中提出了一个基于图表的框架，用于模拟网络攻击对智能电网的物理影响。

2013 年 STPA-Sec 首次被提出[25]，其除了可以评估功能安全以外，还运用 STPA 的基本方法和原理，分析、发现可能出现的信息安全隐患。为整个系统定义控制层，并详细分析每个控制循环及危险场景，并进行组件层上的信息安全评估，寻找最有效的解决措施。该方法将抽象的控制层（注重功能连接、控制概念及算法）映射到组件层（部署控制算法和传感器节点、使用的网络节点、物理网络连接和应用程序级协议等系统实现范畴），此外还把信息安全约束引入到分析中，即在定义系统级约束时同时考虑功能安全和信息安全，最后在一种信息物理系统的实例中进行了相应的危险分析应用。但 STPA-Sec 人为信息安全是直接与功能安全相关的，而文献[2]提出违反机密（欺诈等行为）、信任的丢失等可能导致的货币损失是与功能安全没有直接关系的，这使得功能安全和信息安全的评估不能简单混为一谈。

3 总结

传统的危险分析方法只适用于单个组件的线性系统，由于工业控制系统中出现了越来越多以信息通信技术为基础的应用。日益复杂、互联的 CPS 导致了更多的功能安全和信息安全隐患，因而需要能同时独立地评估两种功能安全和信息安全的方法。由传统的只能评估物理系统的危险分析方法，发展到能评估含信息通信技术的控制系统的功能安全，再到能同时评估这类系统的功能安全和信息安全，控制系统的危险分析方法日益完善。新的危险分析方法——STPA 最初应用

在航空航天上，后来被逐渐应用在智能电网控制系统和汽车领域中，但随着智能化和网联化的发展，汽车安全形势变得更加复杂，功能安全和信息安全相互交融，综合安全保护意义重大，更带来严重挑战，研究如何将 STPA、STPA-Sec 以及 STPA-SafeSec 等危险分析方法应用到智能网联汽车领域十分重要。

参考文献

[1] Mitchell R , Chen I R . Effect of Intrusion Detection and Response on Reliability of Cyber Physical Systems[J]. IEEE Transactions on Reliability, 2013, 62(1):199-210.

[2] Friedberg I , Mclaughlin K , Smith P , et al. STPA-SafeSec: Safety and security analysis for cyber-physical systems[J]. Journal of Information Security and Application, 2017,34:183-196.

[3] Koscher K , Czeskis A , Roesner F , et al. Experimental Security Analysis of a Modern Automobile[C]// 31st IEEE Symposium on Security and Privacy, S&P 2010, 16-19 May 2010, Berleley/Oakland, California, USA. IEEE, 2010.

[4] SAE J3061. Cybersecurity Guidebook for Cyber-Physical Vehicle Systems. https://www.sae.org/standards/content/j3061_201601/.

[5] Haasl D F, Roberts N H, Vesely W E, et al. Fault tree handbook[J]. 1981.

[6] King P . Ten Questions About Human Error, A New View of Human Factors and System Safety - [Book reviews][J]. IEEE Engineering in Medicine and Biology Magazine, 2005, 24(3):18-18.

[7] Manning C . Dekker S. The Field Guide to Understanding Human Error.[J]. Aviation, 2007, volume 78(2):148-148(1).

[8] Lutz R R . Analyzing software requirements errors in safety-critical, embedded systems[C]// Requirements Engineering, 1993. Proceedings of IEEE International Symposium on. IEEE, 1993.

[9] Goetsch, Steve. Safeware: System Safety and Computers, by Nancy Leveson[J]. Medical Physics, 1996, 23(10):1821.

[10] Fleming, Cody Harrison, Spencer, Melissa, Thomas, John,et al. Safety assurance in NextGen and complex transportation systems[J]. Safety Science, 55:173-187.

[11] Ericson C A. Fault tree analysis[C]//System Safety Conference, Orlando, Florida. 1999, 1: 1-9.

[12] Duckworth H A, Moore R A. Social responsibility: Failure mode effects and analysis[M]. CRC Press, 2010.

[13] Takeichi M , Sato Y , Suyama K , et al. Failure rate calculation with priority FTA method for functional safety of complex automotive subsystems[C]// 2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering. IEEE, 2011.

[14] Jordi Dunjó, Fthenakis V , Juan A. V Íchez, et al. Hazard and operability (HAZOP) analysis. A literature review[J]. Journal of Hazardous Materials, 2010, 173(1-3):19-32.

[15] Avizienis A , Laprie J C , Randell B , et al. Basic concepts and taxonomy of dependable and secure computing[J]. IEEE Transactions on Dependable and Secure Computing, 2004, 1(1):11-33.

[16] Leveson N . A New Accident Model for Engineering Safer Systems[J]. Safety Science, 2004, 42(4):237-270.

[17] Nicolas Dulac, Nancy Leveson. An Approach to Design for Safety in Complex Systems[J]. Incose International Symposium, 2004, 14(1):517–530.

[18] Weiss K A, Dulac N, Chiesi S, et al. Engineering spacecraft mission software using a model-based and safety-driven design methodology[J]. Journal of Aerospace Computing, Information, and Communication, 2006, 3(11): 562-586.

[19] Lutz R R . Analyzing software requirements errors in safety-critical, embedded systems[C]// Requirements Engineering, 1993. Proceedings of IEEE International Symposium on. IEEE, 1993.

[20] Thomas IV J P. Extending and automating a systems-theoretic hazard analysis for requirements generation and analysis[D]. Massachusetts Institute of Technology, 2013.

[21] Thomas J, Sgueglia J, Suo D, et al. An integrated approach to requirements development and hazard analysis[R]. SAE Technical Paper, 2015.

[22] Dondossola G , Szanto J , Masera M , et al. Effects of intentional threats to power substation control systems[J]. International Journal of Critical Infrastructures, 2008, 4(1/2):129.

[23] Wang W , Lu Z . Cyber security in the Smart Grid: Survey and challenges[J]. Computer Networks, 2013, 57(5):1344-1371.

[24] Kundur D , Feng X , Mashayekh S , et al. Towards modelling the impact of cyber attacks on a smart grid[J]. International Journal of Security and Networks, 2011, 6(1):2.

[25] Young W, Leveson N. Systems thinking for safety and security[C]//Proceedings of the 29th Annual Computer Security Applications Conference. ACM, 2013: 1-8.