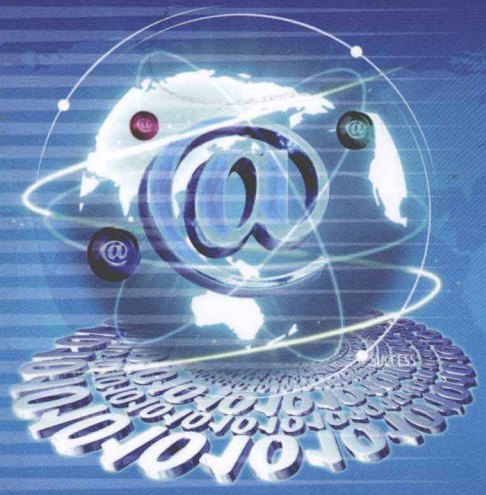
网络信息系统可生存性技术研究

张乐君 著



HEUP 哈尔滨工程大学出版社
Harbin Engineering University Press

网络信息系统可生存性技术研究

张乐君 著

HEUP 哈尔滨工程大学出版社
Harbin Engineering University Press

内容简介

计算机网络与信息系统广泛地应用于工业、商业、政府和国防部门。网络系统内、外部环境日趋复杂使得任何网络系统都不可能是安全的。因此,信息系统的可生存性技术研究是一个重要的研究方向。本书所做的主要工作是研究网络信息系统的可生存性评估和增强技术,达到能够分析和定位信息系统可生存性弱点,并在给出量化分析结果的基础上,提出改进可生存性状况的增强方法。

本书可作为网络与信息安全专业硕士研究生或博士研究生的教学参考用书,也可作为从事网络系统生存性技术研究人员的参考用书。

图书在版编目(CIP)数据

网络信息系统可生存性技术研究/张乐君著. —哈尔滨:
哈尔滨工程大学出版社,2012.3
ISBN 978-7-5661-0326-0

Ⅰ.①网… Ⅱ.①张… Ⅲ.①计算机网络-信息系统-
系统安全性-研究 Ⅳ.①TP393.08

中国版本图书馆 CIP 数据核字(2012)第 037530 号

出版发行	哈尔滨工程大学出版社
社　　址	哈尔滨市南岗区东大直街 124 号
邮政编码	150001
发行电话	0451-82519328
传　　真	0451-82519699
经　　销	新华书店
印　　刷	黑龙江省地质测绘印制中心
开　　本	787mm×960mm 1/16
印　　张	8.25
字　　数	125 千字
版　　次	2012 年 3 月第 1 版
印　　次	2012 年 3 月第 1 次印刷
定　　价	20.00 元

http://press.hrbeu.edu.cn
E-mail:heupress@hrbeu.edu.cn

前　言

随着国家信息化带动工业化发展战略的确立，计算机网络与信息系统建设取得了长足的发展，网络系统广泛地应用于工业、商业、政府和国防部门。网络系统内、外部环境日趋复杂使得任何网络系统都不可能是安全的，因此，信息系统的可生存性技术研究是继系统安全性之后的又一个重要研究方向。

本书所做的主要工作是研究网络信息系统的可生存性评估和增强技术，达到能够分析和定位信息系统可生存性弱点，并在给出量化分析结果的基础上，提出改进可生存性状况的增强方法。本书具体包括以下内容。

第一，利用随机 Petri 网模型对信息系统可生存性分析进行了建模研究。第一步，将信息系统抽象为请求组件、通信组件、处理组件和存储组件四个部分；第二步，将信息系统工作流程形式化描述和可生存性分析建模相结合，分别描述了通用信息系统、系统组件失效修复、串联并接、冗余以及具有可生存属性组件的建模方法；第三步，通过模拟的方法实现对信息系统形式化描述的同时对其生存能力作了定性和定量分析。

第二，提出了一种信息系统可生存性的层次化评估模型。该模型根据系统中关键服务及其原子服务组件信息进行漏洞探测，生成网络攻击图和攻击方案，并对系统进行基于真实环境下的可生存性测试，记录攻击中和攻击后的服务质量变化。通过服务质量变化分析系统可生存性的四个关键属性，最终实现对整个网络系统的可生存性量化分析。

第三，针对分布式系统的特点，提出了一种基于测度属性关系网络划分的生存性异常检测方法，并以此为基础建立了网络模型来刻画分布式系统测度属性的关系；进而，提出了基于隐马尔可夫模型的测度属性关系网络划分方法，并利用最新监测的属性关联度信息，设计了基于测度属性关系网络划分的分布式系统异常检测算法；通过对 DoS，U2R 和 PROBE 三种攻击进行异常检测实验，从而验证了所提出方法的有效性。

第四，在系统冗余和多样性的前提条件下，提出了原子组件自组织来增强系统可生存性方法；根据中心极限定理以及系统运行历史数据，获取原子组件可生存性变化曲线，并结合系统服务效率以及服务质量等因素，设计了基于自组织的可生存性增强算法。

第五，在介绍了多种连接迁移技术的基础上，设计一个基于连接迁移技术的服务自组织系统。该系统根据服务处理流程将系统组件分解为通信组件、服务分发组件、服务提供组件；备份组件将其可生存性信息实时发送给工作组件，并获取最新服务状态列表，当某一组件可生存性最高时，根据组件的功能及其在体系结构中的位置，通过多种连接迁移技术进行服务自组织。该系统具有对用户透明，配置简单灵活的优点。实验证明，该方法可以有效提高服务质量，并达到增强服务可生存性的目的。

本书得到了国家自然科学基金青年科学基金（61100008）、黑龙江省自然科学基金面上项目（F201023）、黑龙江省教育厅科学研究基金资助项目（11553045）以及中央高校基本科研业务费（HEUCF100602）的支持，在此表示感谢。

本书在写作过程中，哈尔滨工程大学计算机科学与技术学院的杨永田、张健沛、赵春晖、杨静、国林等老师，对本书的研究结构和研究内容提出了许多宝贵的意见，在此深表谢意。同时，在本课题研究与本书撰稿过程中参考或引用了国内外许多专家学者的论著，对他们的辛勤劳动作者表示衷心感谢。另外，还要感谢哈尔滨工程大学出版社老师们的辛勤劳动，使得本书能够顺利出版。

网络系统生存性技术的研究正越来越受到学术界的关注，本书虽尽力将作者所做的研究成果展现出来，但在研究过程中难免存在不足之处，我们将在后续的研究中进一步完善与升华。同时，也恳请读者批评指正，共同促进网络系统可生存性研究技术的进一步发展。

著　者

2012 年 1 月

目　录

第1章 绪 论

1.1 研究背景

计算机网络技术是带动整个人类社会科技进步的一个强劲动因。随着基于网络的信息系统在社会各部门中日益广泛地应用,其作用也逐步增大。以网络为手段获得信息和交流信息已经成为现代社会的一个重要特征。总之,网络信息系统已经成为人类社会生活中不可或缺的组成部分。网络系统内、外部环境日趋复杂使得任何网络系统都不可能是安全的,并且由于网络系统中存在着大量不稳定因素,网络信息系统的可生存性已逐步成为科技界关注的焦点。

网络安全技术发展到今天一共经历了三个阶段。

第一阶段:入侵阻止。关键思想是"防",采用加密、认证、安全分级、访问控制等办法来保证信息的安全,阻止入侵。

第二阶段:入侵检测。关键思想是"检",网络入侵防不胜防,所以要对不能防的入侵行为进行检测,于是出现了防火墙、IDS 等入侵检测技术。

第三阶段:可生存性研究。可生存性的主要思想可以总结为"容",这也是目前网络安全研究的新的热门方向。

网络系统的可生存性正是在这个阶段提出来的。由于网络系统可生存性的研究与传统的网络安全理念不同,它更多地关注持续服务的能力,因此能源、金融、政府等相关部门对其关注程度越来越高。从国家层面上来说,该领域的研究也得到了充分的重视,针对国家基础设施的关键服务网络系统可生存性的研究也在如火如荼的开展。

信息系统可生存性定义是 Barnes 等人于 1993 年提出的,其是指系统提供关键服务的能力,即系统在面临攻击、失效和偶然事件的情况下仍然可以按照需求及时

完成任务的能力。在完成关键服务的同时系统仍然保持其基本属性，如数据完整性、机密性等及其他属性。可生存性是衡量网络系统“容侵、容错、容灾”能力的重要依据。网络系统的可生存性是建立在传统安全理论、可信计算等基础之上，其正在成为一个新的研究方向。

网络系统可生存性概念的提出得到广泛的重视，究其原因在于网络系统可生存性概念具有丰富的内涵和不同的安全理念。首先，网络安全追求目标的转移。可生存性基本含义表明，系统安全目标不再是保证整个系统的绝对安全可靠，而是保证系统中提供的一些关键服务的正确性、连续性、无间隙性。例如，一个由服务器集群构成的服务网络，其可生存性并不关心哪台服务器是否出现了故障，而是关心在部分服务器出现故障的时候，服务网络是否还能继续提供服务。其次，不存在绝对的可生存性。可生存性研究的一个关键假设是系统的任何部件均可能发生故障或被入侵，然而，没有一个系统能够应对不受限制的恶意故障或攻击。因此，一个系统的可生存性总是相对于给定的内外部环境（故障模型和攻击模型）而言的，绝对的可生存性实际上是不存在的。再次，可生存性含义具有多样性。可生存性的物理含义与具体系统提供的服务内容是密切相关联的，如通信网络的可生存性，通常指网络在节点或通信链路失效后的联通能力；网络文件系统的可生存性指部分存储器因随机故障或恶意攻击导致 Byzantine 失效后，仍可以提供正确的、一致的数据访问能力。因此，通用的、形式化的定义是可生存性研究的一个基本问题。但是，由于可生存性概念包含了跨多种学科的内容，同时由于研究对象的广泛性，到目前为止，学术界尚未取得一致的定义，也没有形成统一的、公认的度量标准，但这方面的努力仍在继续。最后，可生存性是网络系统的整体特性。系统整体的可生存性设计得好，有可能比某些组成部分的可生存性强；某些功能甚至可能不具备可生存性，但整个系统仍有良好的可生存性。

综上所述，可生存性研究通过对特定的入侵模式提供相应的可生存性策略，从而保证系统能够完成既定的需求，提供所要求的正常服务。这些策略主要包括对入侵集的抵抗、识别、应急响应和快速恢复。在系统遭受攻击时，通过可生存性策略保障关键功能服务集合前提下，尽量恢复其他服务的可生存性。

从下列事件可以了解可生存性研究的重要性。

1993 年，美国世贸中心大楼发生爆炸。爆炸前，约有 350 家企业在该楼中工

作。1 年后,再回到世贸大楼的公司只有 150 家,约有 200 家企业由于无法存取重要的信息系统而倒闭、消失。

1995 年 1 月,日本神户地区大地震摧毁了约 1 700 部电脑系统,造成约 1 000 多亿美元的损失。

1996 年 1 月,美国南加州洛杉矶 6.6 级地震损坏的计算机系统和网络,造成约 300 亿美元的损失。

1999 年 6 月,美国一家著名的商业交易网站的主机宕机,由于 24 小时内未能恢复访问,事件发生的两个星期后,该公司的股票市值下跌了 36%。

我国台湾地区,在 1999 年大地震之前,各公司对灾备都不重视,地震发生后,由于关键业务中断为公司带来了损失,很多公司才因此认识到灾难恢复非常重要。

2001 年"911 事件"中,由于有 1993 年爆炸的前车之鉴,在美国世贸大楼的多数公司都建起了自己的灾备系统,因此当灾难再次降临时,有一批公司仍可及时地通过自己的数据恢复计划来重整旗鼓;而另外一些企业,则在爆炸发生后因丢失了关键业务数据而倒闭。

2003 年,国内某电信运营商的计费存储系统发生两个小时的故障,造成 400 多万元的损失,这些还不包括导致的无形资产损失。

据 IDC 的统计数据表明,美国在 2000 年以前的 10 年间发生过灾难的公司中,有 55% 当时倒闭,剩下的 45% 中,因为数据丢失,有 29% 也在两年之内倒闭,生存下来的仅占 16%。Gartner Group 的数据也表明,在经历大型灾难而导致系统停运的公司中有 2/5 再也没有恢复运营,剩下的公司中也有 1/3 在两年内破产。

根据有关机构统计,对关键业务运行要求最高的是银行业,每次计算机系统宕机导致的损失平均为 1 000 万美元,同时还会导致无法估量的无形资产损失,而通过采取灾难数据的恢复方案总共花费平均只有 100 万美元。

在信息系统可生存性的研究中,主要包括可生存性分析和可生存性增强两个领域,在本书中将分别对其进行研究。所谓可生存性分析就是对系统的可生存性进行评估和分析,给出系统的可生存性状况并提出具体的改进建议,以便指导系统的可生存性设计。需要进行可生存性分析研究的一个重要原因是确定影响可生存性的关键点,通过分析系统的可生存性需求为可生存性增强提供帮助和指导。可生存性分析的最终目标是可利用具体数据对其进行较为准确的描述,实现对系统

可生存性的量化分析。可生存性增强是指采用某种手段,提高系统服务能力和质量,改进系统的可生存性状况的方法。可生存性增强是可生存性分析的最终目的,可生存性分析为可生存性增强提供依据和支持。

1.2 研究现状

对于可生存性技术的研究,最早出现在传统领域,在第一次世界大战和第二次世界大战的时候已经在军事设备上进行了探索,如海军舰艇在遭受攻击时的抗沉没能力,飞机在某些部件失效的时候如何继续飞行并保证飞行员的生存;在民用领域中也进行了如建筑物的防火能力等的研究。

进入网络时代后,人们也开始关注网络通信系统的可生存性问题,因为网络系统与传统领域中可生存性研究对象性质的不同,出现了一系列的技术挑战。针对通信系统的可生存性研究,主要是从硬件上对通信设备的性能考虑系统的生存能力,而且一般是从网络链路的连通性来分析的,这是目前可生存性研究较多的领域。国际组织 TIA1.2 工作组负责发展和推荐主要针对电信网络可生存性设计和测评的标准和技术报告。相对于美国和欧洲的研究而言,我国的可生存性技术研究开展得较晚、较少,这也是由于我国安全领域研究起步较晚等一系列因素影响的结果,直到 1999 年我国的应急组织 CCERT 才成立。目前国内的研究工作正针对 ATM,SDH 等一些特定的通信网络进行可生存性的分析和设计。

针对信息系统的可生存性研究,相对而言起步较晚。借助于 Vickie R. Westmark 对 IEEE,ACM 和 SEI 上相关论文进行的资料统计,结果见表 1.1。

表 1.1 可生存性研究论文统计

检索方式	First Pass				Second Pass			
	ACM	IEEE	SEI	总计	ACM	IEEE	SEI	总计
主题词	1 370	1 799	203	3 372	21	210	11	242
作者	83	208	97	388	6	20	2	28
总计				3 760				270

检索使用的主题词包括可生存性、可生存的、基于构件的软件工程、基于构件的分布式系统、分布式网络系统、分布式网络环境；作者包括 Knight, Sullivan, Yacoub, Weiss, Hofman, Ellison。First Pass 是指只要题目与分布式环境下的可生存性相关即可。Second Pass 是指对通过 First Pass 的论文分析其摘要、前言、结论等，与可生存性定义、测量和实现相关的论文的统计结果。

从 First Pass 到 Second Pass 文章数量的巨大落差(从 3 760 篇下降到 270 篇)可以看出，目前的研究还不成熟。这 270 篇论文可以归结为以下 7 个分类：

(1)仅仅提到可生存性，根本没有相关研究；

(2)仅仅提供可生存性的定义；

(3)认识到可生存性的重要性，但是没做实际工作；

(4)认识到可生存性测定标准的重要性，但是没做实际工作；

(5)认识到可生存性实现的重要性，但没做实际工作；

(6)提供可生存性实现的方法，但是很少或者几乎没有在实际应用中使用；

(7)很好地实现了可生存性。

这些论文绝大多数都归于前 5 个分类，仅仅是认识到可生存性的重要性，还缺乏对可生存性的统一认识和标准，论文中提及的可生存性实现也都基于非正式应用，还没有经过实际的应用的检验。最后一类很好地实现了可生存性，到目前为止还只是美好的展望和期待。

目前可生存性研究已经成为国内外网络安全研究领域的热门话题，开展的相关研究很多，但大都围绕着可生存性是什么、如何测定可生存性以及如何实现可生存性等三个问题进行的，并且形成了各自的认识和理解。下面本书将分别对以上三个问题的国内外典型研究成果进行介绍。

1.3 可生存性定义

可生存性其实不是一个新的概念名词，在军事和电信领域早已出现并使用。到目前为止，学术界尚未对网络信息系统的可生存性取得一致的定义，也没有形成统一的、公认的度量标准，但这方面的努力仍在继续。

Deutsch 从软件工程的角度提出可生存性的定义：在系统部分瘫痪的情况下，关键服务还能够使用的程度。

IEEE 定义的可生存性是指系统一部分无法工作时，软件系统能够无故障地执行和维持关键功能的能力。

以 Ellison 等为代表的 CMU/SEI 的研究小组则提出了下面的定义：可生存性是指在遭受攻击、故障或意外事故时，系统能够及时完成其关键任务的能力。

上面仅仅列出了几个典型定义，目前各种论文和文献中提及的可生存性定义有近 50 个。这些定义没有任何一个能够得到学界和业界的普遍认可，包括 IEEE 给出的定义，其中影响相对较大的还属于 CMU/SEI 的研究小组给出的定义。

目前可生存性还没有统一的严格定义，但是定义可生存性时，本书认为要考虑以下 5 个因素。

(1)系统：包括系统的体系结构，关键服务类型，定义系统任务和主要功能，以及系统是有边界系统还是无边界系统等。

(2)环境：定义系统使用环境，要明确提出定义可生存性的网络环境。

(3)威胁：可能影响到系统提供服务的威胁因素，主要分为故障、入侵和灾难。

(4)服务持续性：服务持续性应该作为系统需求进行定义，系统及网络性能的下降不应该被用户所觉察。

(5)响应时间：服务应该在系统要求或者用户所期望的时间内可用，并能够正确作出回应。

1.4 可生存性评估模型

对于给定的系统如何评估其生存能力，已有的研究表明：对于不同的对象，由于建立系统模型机理的区别，采用的方法也不完全相同。各类通信网络的可生存性分析评估通常建立基于图的模型。攻击或故障模拟为图的节点或边的切除，其可生存性分析多数通过计算图的某种连通能力来度量，已经取得较多的成果。

下面将分别从系统结构、服务组件、数据流图、潜在攻击、系统与环境关系等方面介绍典型的评估模型。

1.4.1 基于系统结构的评估模型

网络信息系统的硬件物理设施通常可以表示为图的某种形式，而服务流可以抽象为调度问题，因而可以用图对系统网络拓扑或物理结构进行建模，进而分析系统的可生存性。在利用这种模型进行分析时，一般利用网络节点之间的连接情况或者一些网络性能参数，如网络流量等。卡内基梅隆大学的 SEI 研究中心提出了可生存性分析方法（Survivable Network Analysis，SNA）可生存性分析方法，从而达到提高系统在受到威胁时的生存能力。

SNA 可以在系统的生命周期、需求分析以及体系结构三个层次进行，最终提出关于分析结果和建议的报告，分析结果通过可生存性图的形式总结出来，列举了当前和推荐的体系结构的策略。Krings 提出一个四步模型，将可生存性分析转化为一个参数化的图模型，该模型是由图和调度算法组成的分析方法的基础，如图 1.1 所示。

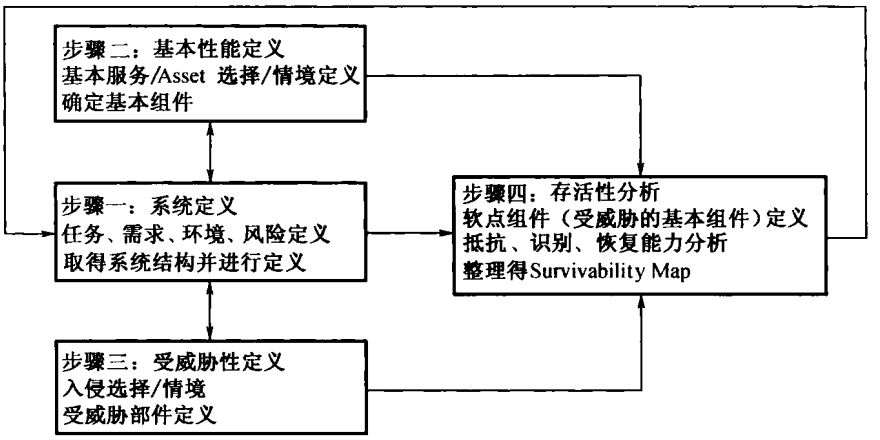


图 1.1 SNA 分析方法

（1）系统定义：主要任务是定义系统任务和主要功能需求；根据系统用户能力和所在区域、系统事务的类型和规模，定义系统使用环境；根据可能遇到不利条件的类型回顾系统风险；根据系统硬件组件和连接方式、软件配置和驻留信息确定体

系结构等。

(2)基本性能定义:这一阶段的主要任务是从系统提供的服务和数据中选择系统所要提供的关键服务和数据。关键服务和数据分别是系统在遭遇入侵、故障或意外情况下,也必须提供的服务和数据。

(3)系统受威胁性定义:这个阶段要求根据操作系统的操作环境选择具有代表性的入侵方法,定义入侵使用情境,进行跟踪确定受威胁的部件。

(4)存活性分析:软点(Soft Spot)组件是指受威胁的基本组件。基本组件是指支持系统关键服务和数据的组件,如果基本组件遭到破坏,系统就无法提供应有的关键服务,可生存性就自然受到影响。所以 SNA 的最后一个步骤就是要对系统的软点组件进行分析。通过对系统软点组件的 3R(Resistance,Recognition,Recovery)特性的分析,得到 SNA 的分析结果,提出体系结构级的相应修改建议。

大量的研究工作都以 SNA 分析方法为基础展开,而传统的 SNA 分析方法只能对系统的 3R 性能作出"定性"的分析。高献伟等人在 SNA 分析的基础上,分别对入侵的危险程度、识别率、服务恢复时间等性能确定了参数模型,对 3R 的各个性能提出一个"量化"方案。

总体来说,SNA 还处于探索阶段,还没有形成规范化的规则集用于分析;在 3R 量化分析中,各种参数的选择需要大量成功案例支持;由于网络系统攻击技术不断的发展,这些策略和案例需要频繁的检查和不断的改进。

1.4.2 基于系统服务组件的评估模型

该模型以信息系统提供的服务组件为中心,根据服务涉及到的系统组件将系统进行简化和建模,在数学上通常表现为一个类似树的结构,该模型把系统的可生存性评估分解为对各个组件可生存性的计算,把系统整体可生存性分解为各个组件可生存性问题来求解,通过综合评估所涉及到的各个组件的可生存性情况,得到该系统的可生存性状态。

郭博渊等人将系统服务的故障归结到原子服务(某个硬件、软件或它们的组合)中去,简化了可生存性的计算,通过对分布式网络系统中系统与配置、服务之间支持依赖关系的定量描述来刻画系统服务的可生存性。DSO National Laboratories 将系统可生存性量化任务分解为四个层次,通过计算各服务组件可生存性,最终归

结出系统的可生存性。

因为该模型从系统组件出发,所以要详细了解系统各个组件的设计方法及其层次结构,以便对系统服务组件进行充分评测;其次,组件可生存性与系统可生存之间的依赖关系缺乏通用的规范化标准。受网络规模的限制,如何从真实系统体系结构自动生成评估模型还是一个难点。

1.4.3 基于数据流图的评估模型

包秀国、陈庆家等人提出了基于生存数据流图的评估方法。该方法的核心思想是:采用问题空间转换思路,基于数据流关系建立系统的一种图模型,其可生存性计算转换为图的连通性计算。首先,将配置操作指令的执行过程看成是由消息内容的分解、传递、转换、细化等一系列对数据流操作处理构成的;将与可生存性相关的所有元素的功能统一抽象为数据流的输入、处理和输出三个环节,构造一种反映元素之间数据流逻辑关系的图模型,将可生存性计算转换为图的连通性计算。采用 Monte Carlo 方法为计算系统可生存性构造一个合适的概率模型,依照模型进行大量的统计实验,然后通过模型或过程的抽样实验来计算所求参数,最后给出可生存性的近似值。

该模型是一种对系统可生存性评估的模拟实验方法,基于服务组件的评估模型注重服务节点,基于数据流图模型关注服务逻辑连接图的连通性,是从另一个角度刻画可生存性。此外,该模型亦可应用到实际网络中,根据数据流的响应时间、丢包率等情况将数据链路连通情况划分为多个等级,以此为根据量化系统在实际环境中的可生存性。Monte Carlo 方法可用于分析十分复杂的对象、复合故障和攻击的影响,并可从多角度进行分析,其主要缺点是该算法需要消耗大量计算资源。

1.4.4 基于潜在攻击的评估模型

该模型的核心思想是通过考察系统服务在不同级别攻击下的服务质量来量化网络系统的可生存性,响应时间是衡量服务质量的重要指标。

McDermott 首先定义系统服务状态(正常状态、攻击响应状态、报警状态、失效状态),根据网络攻击的随机性,采用随机进程代数(Stochastic Process Algebra, SPA)模拟攻击者和系统的行为,并建立攻击过程模型;其次,用转换概率矩阵建立

系统面临攻击的反应模型;最后,采用数学公式分析网络系统遭受攻击后系统的终结状态来量化网络系统可生存性。这种研究方法是基于预测(即潜在攻击下)的可生存性,因此攻击者的能力和持续攻击的可能性也应考虑在内,极大地增加了计算复杂性,DoS 攻击下网络系统服务的状态迁移情况如图 1.2 所示。

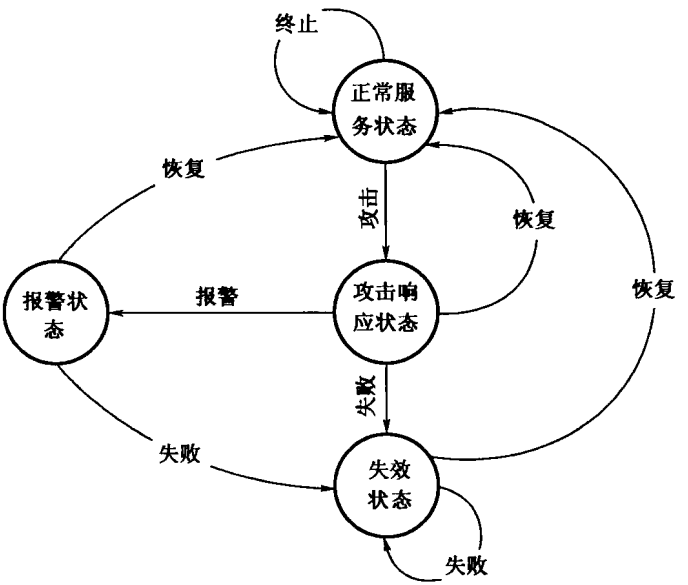


图 1.2 DoS 攻击下网络系统状态迁移图

此外,也可利用真实攻击对网络系统可生存性进行测试分析。通过考察真实攻击对网络系统的影响量化系统的可生存性。但某些攻击会对网络系统产生不可恢复影响,而且攻击的复杂性和多样性不断增加、发展,因此缺乏一种有效的攻击探测和分析手段。另外,该方法未考虑网络系统外部环境不断变化等多方面因素影响,导致这种方法研究存在一定的局限性。

1.4.5 基于系统与环境关系的评估模型

该模型认为:从系统科学的角度来看,信息系统可视为一个开放的复杂巨系