

基于容斥原理的网络可靠度分析与计算

李志豪¹

(1. 西安电子科技大学 计算机科学与技术学院, 陕西 西安 710071)

摘要: 网络的可靠性是现代网络设计、规划及建设必须考虑的因素,而如何去评估网络的可靠度也成为最近研究的重点。在万维网中,数据报的发送与接受,必须保证用户、设备、系统之间的信息准确、及时无误地传输,其中可靠度就是网络传输可靠性中的一个重要指标。本文通过组合数学中容斥原理,利用传统的解析算法探究了通信网络在数据传输中可靠性的评估,对网络的可靠度进行了分析与量化,并提出了一种高效的求解算法。在实际应用场景中,本文以高校实验机房的计算机网络系统为例进行着重分析计算,结合实际情况验证了理论的正确性;通过实例初步发现网络可靠度与网络设备失效概率分布和网络链路失效概率分布存在紧密联系,因此本文通过计算机仿真实验,模拟探究了网络可靠度与网络设备失效概率分布和网络链路失效概率分布之间的联系。

关键词: 网络可靠度; 容斥原理; 网络设备失效概率分布; 网络链路失效概率分布

中图分类号: 10-19665 **文献标识码:** A **文章编号:** 1001-2400-2023-12-17

Network Reliability Analysis and Calculation Based on the Inclusion-exclusion Principle

LI Zhihao¹

(1. School of Computer Science and Technology, Xidian University, Xi'an 710071, China)

Abstract: Network reliability is a factor that must be considered in modern network design, planning and construction, and how to evaluate network reliability has become the focus of recent research. In the World Wide Web, the sending and receiving of datagrams must ensure accurate, timely and error-free transmission of information between users, devices, and systems. Reliability is an important indicator of network transmission reliability. This paper uses the principle of inclusion and exclusion in combinatorial mathematics and uses traditional analytical algorithms to explore the reliability evaluation of communication networks in data transmission, analyzes and quantifies the reliability of the network, and proposes an efficient solution algorithm. In practical application scenarios, this article takes the computer network system in the experimental computer room of a university as an example to conduct focused analysis and calculation, and verifies the correctness of the theory based on the actual situation; through examples, it is initially discovered that network reliability, network equipment failure probability distribution and network link failure Probability distributions are closely related, so this paper uses computer simulation experiments to simulate and explore the connection between network reliability and network equipment failure probability distribution and network link failure probability distribution.

Key Words: Network Reliability; Inclusion-exclusion Principle; Network Equipment Failure Probability Distribution; Network Link Failure Probability Distribution

收稿日期: 2023-12-18

网络出版时间: 2023-12-19

基金项目: 国家自然科学基金 (61931018)

作者简介: 李志豪 (2003—), 男, 西安电子科技大学本科生, E-mail: lizhihaoms@xidian.edu.cn

通信作者: 纪建 (1988—), 女, 教授, E-mail: jijian@mail.xidian.edu.cn

网络出版地址: <https://doi.org/10.19665/j.issn.1001-2400.2023.12.17>

1 引言

随着现代网络技术的飞速发展,网络已经渗透到人们生活的各个方面,使得人们对网络的依赖度越来越高,网络承载着越来越多、越来越重要的业务。然而网络故障经常会导致严重的后果,例如网络流量负载拥堵使得系统崩溃,用户数据库发生故障导致业务流失等,因此在利用网络开展各种业务时,我们更需要考虑所依赖网络的可靠性问题。网络的可靠性是现代网络设计、规划及建设必须考虑的因素,而如何去评估网络的可靠度也成为最近研究的重点。在万维网中,数据报的发送与接受,必须保证用户、设备、系统之间的信息准确、及时无误地传输,其中可靠度就是网络传输可靠性中的一个重要指标。

但网络的复杂性、动态性、多态性等特点使得传统成熟的可靠性评估方法,如可靠性框图(RBD)法、故障树分析(FTA)法等很难用于网络。近 10 年来,对网络可靠性的研究取得了很多成果,网络可靠性的内涵也由传统的基于网络拓扑结构的连通可靠性逐渐拓展到考虑网络流的容量可靠性,并向基于业务等考虑用户需求的性能可靠性延伸。

在本文中,我们着重研究基于网络拓扑结构的连通可靠性,利用组合数学中的容斥原理提出一种新的评估体系,建立网络传输的可靠性指标,利用现有数据进行指标的量化与评估模型的构建,最后通过计算机软件仿真,在虚拟环境中随机模拟网络传输情形,开展实验进行验证。

2 网络连通可靠性的评估

2.1 问题的表述

传统方法对网络的建模,可以表示为将网络抽象成由一组节点集与一组链路集构成的图 $G(V, E)$,

$$V = \{v_1, v_2, \dots, v_n\}, E = \{e_1, e_2, \dots, e_n\} \quad (1)$$

1955 年, Lee 在“Analysis of Switching Networks”(交换网络分析)一文中,定义了以“能实现连通功能的概率”为度量的端可靠度,首次使用了以连通为规定功能的可靠性指标[1]。连通可靠性也是最早提出的网络可靠性指标,研究将网络按照有无指定的源点分为有源网络与无源网络。由于有源网络在实际应用中更加常见,因此在本文中,主要针对有源网络的可靠性进行探讨。

有源网络可靠性指标分为 3 类:

- ST 可靠度: 源点 s 与终点 t 保持连通的概率;
- SK 可靠度: 源点 s 与特定的端点集 K ($K \subset V$, V 是网络中所有端点的集合)保持连通的概率;
- SAT 可靠度: 源点 s 与网络中所有其他端点保持连通的概率;

在本文的研究中,针对有源网络传输,网络可靠度问题主要包括以下两个子问题:

- 1) ST 可靠度问题,即计算信息从源点 s 能够可靠地传输到终点 t 的概率;
- 2) SK 可靠度问题,即计算信息从源点 s 能够传输到特定终点集合 K 中所有节点的概率;

2.2 问题的假设

针对网络可靠性评估提出的两个问题,利用经典解析算法需要对所研究问题做出以下两个假设,以保证建立模型的合理性与可解释性。

- 网络中的链路只有故障、正常两种状态;
- 网络中链路故障的概率是统计独立的;

2.3 符号约定

表 1 符号约定说明

符号	说明
$G(N, E)$	节点集为 N , 边集为 E 的有向网络
$P(s, t)$	从源点 s 到终点 t 能够通信的概率
\bar{N}_i	\bar{N}_i 是 N_i 的补集, 且 $N_i, \bar{N}_i \subset N, s \in N_i, t \in \bar{N}_i$
(N_i, \bar{N}_i)	网络 G 的极小割, 也记为 C_i

$C_i C_j$	C_i 和 C_j 中的所有边的并集
$\sigma(N_i)$	$\{N_j N_j \supset N_i, \text{按 } N_j \text{ 递增的顺序排列}\}$
r_i	$\sigma(N_i)$ 中含有割集的元素个数
M_{ij}	$\sigma(N_i)$ 中的第 j 个元素
U_i	网络 G 中 i 个割集的并集
ε_i	如果 U_i 中含有割集的数为偶数, 取值 -1 否则取值 $+1$
$R(N_i, N_j) \neq 0$	表示 N_i 中的每个节点与 N_j 中的每个节点都连通
$R(N_i, N_j) = 0$	表示 N_i 中的每个节点与 N_j 中的每个节点都不连通
$N^+(v_i)$	集合 $\{v_j \text{有向边 } (v_i, v_j) \in E\}$

2.4 模型的建立与求解

在经典解析算法中, 利用组合数学中的容斥原理是一种常见的方法, 这种方法可以形式化地定义为:

容斥原理法: 将网络可靠度表述为全部最小路集的并 (或将网络不可靠度表示为全部最小割集的并), 然后采用容斥原理去掉相容时间相交的部分, 计算相应的可靠度。

在第一个子问题即两终端可靠度问题中, 两终端可靠度的容斥原理方法就是找出所有的道路或割集, 然后代入容斥原理公式中求得可靠度。计算复杂度分析: 如果直接应用容斥原理计算, 表达式中的项数随割集或道路的增加呈指数增长, 计算量很大, 但表达式中会有很多相互抵消的项, 因此找出容斥原理公式中的相消项或不消项是解决网络可靠度的有效方法。

在无圈有向网络 G 中, 假设节点总是正常工作的且每条边只有正常工作或失效两种状态, 每条边的工作状态是相互独立的, 且 C_1, C_2, \dots, C_m 是 G 的 m 个极小割, 利用容斥原理公式得 s, t 两终端可靠度为

$$P(s, t) = 1 - Q(s, t) \quad (2)$$

其中,

$$\begin{aligned} Q(s, t) &= P_r(C_1 \cup C_2 \cup \dots \cup C_m) \\ &= \sum_{i=1}^m P_r(C_i) - \sum_{i < j} P_r(C_i C_j) + \dots + (-1)^{m-1} P_r(C_1 C_2 \dots C_m) \end{aligned} \quad (3)$$

在 $Q(s, t)$ 的表达式中共有 $2^m - 1$ 项, 对于不同的 i, j , 可能有 $U_i = U_j, \varepsilon_i = -\varepsilon_j$, 这样就可以保证 $\varepsilon_i P_r(U_i)$ 与 $\varepsilon_j P_r(U_j)$ 相抵消, 从而减少网络可靠度的计算开销。

a) 网络可靠度的简化计算

在利用容斥原理求解网络可靠度时, 上文已经论述能够通过 $\varepsilon_i P_r(U_i)$ 与 $\varepsilon_j P_r(U_j)$ 相抵消来减少网络可靠度的计算开销。因此根据经典解析算法[2]提出以下两个引理具体阐述抵消项需满足的条件:

引理 1 设 $C_i = (N_i, \bar{N}_i), C_j = (N_j, \bar{N}_j)$ 为网络的两个极小割, N_i, N_j 不能比较, 则在容斥原理公式中同时包含 C_i, C_j 的项被消掉。

引理 2 设 $G(N, E)$ 是一个 $k(k \geq 1)$ 连通的无圈有向网络, $V = (v_1, v_2, \dots, v_n)(n \geq k)$ 是 G 的一个极小点割, N_1, N_1' 是两个节点集合, $s \in N_1, t \in N_1', R(s, N_1) \neq 0, R(t, N_1') \neq 0, V \cup N_1 \cup N_1' = N$, $N_2 = V \cup N_1, N_3 = N_1 \cup \{v_i\}(1 \leq i \leq n)$ 。则 $C_1 = (N_1, \bar{N}_1), C_2 = (N_2, \bar{N}_2), C_3 = (N_3, \bar{N}_3)$ 都是 G 的极小割, 在容斥原理公式中同时包含 C_1 和 C_2 的项被消掉。

b) 网络可靠度的求解算法

根据以上建立的模型, 借助图论知识设计求解一个无圈有向网络的可靠度算法[3]如下:

输入: 网络 G 的所有极小割 C_1, C_2, \dots, C_m 及与之对应的结点集合 N_1, N_2, \dots, N_m , 这里不妨

假设 $|N_1| \leq |N_2| \leq \dots \leq |N_m|$

输出: 两终端可靠度 $P(s, t) = 1 - Q(s, t)$

算法执行:

1. 对于每个节点 i , 寻找对应的集合 $\sigma(N_i)(i = 1, 2, \dots, m)$;
2. 利用子程序 ProduceTree 生成根有向树, 即产生 $Q(s, t)$ 中的所有非相消项;
3. 除根点外, 将根有向树 T 的所有顶点的权相加即得 $Q(s, t)$ 的符号表达式, 最后得网络可靠度的表达式: $P(s, t) = 1 - Q(s, t)$

3 对真实应用场景的分析

3.1 应用场景的表述

某高校实验机房是由多个服务器和交换机组成的计算机网络, 网络中的设备通过冗余连接相互连接。服务器和交换机都有一定的可靠度, 而连接它们的链路也有可能失效。我们希望使用容斥原理法计算整个网络的可靠度, 以评估实验机房网络连通的可靠性。

1) 问题描述

实验机房的网络是一个包含 $n = 6$ 个服务器 (标记为 S_1, S_2, \dots, S_n) 和 $m = 8$ 个交换机 (标记为 SW_1, SW_2, \dots, SW_m) 的计算机网络。服务器和交换机通过链路相连, 每条链路有一定的失效概率。我们希望计算整个网络从某个源点到某个目标点的可靠度。

2) 网络结构

- 每个服务器和交换机都有独立的失效概率, 分别为:

表 2 网络结构参数说明

服务器失效概率	$ps_1 = 0.10$	$ps_2 = 0.15$	$ps_3 = 0.20$	$ps_4 = 0.12$
	$ps_5 = 0.08$	$ps_6 = 0.18$		
交换机失效概率	$psw_1 = 0.10$	$psw_2 = 0.15$	$psw_3 = 0.20$	$psw_4 = 0.12$
	$psw_5 = 0.08$	$psw_6 = 0.18$	$psw_7 = 0.10$	$psw_8 = 0.15$

- 服务器和交换机之间的连接通过链路实现, 每条链路的失效概率为 $p_{link} = 0.05$

3.2 实验结果的分析

从实验机房网络结构中可以看出, 由于每个服务器以及交换机均存在失效的概率, 可以预测到整个网络的可靠度将不会太高, 现利用设计的算法进行求解从源点 S_1 到目标点 SW_8 的可靠度:

先根据网络结构参数求得 $Q(S_1, SW_8)$:

$$Q(S_1, SW_8) = \sum_{i=1}^{m+n} P_r(C_i) - \sum_{i < j \leq m+n} P_r(C_i C_j) + \dots + (-1)^{m+n-1} P_r(C_1 C_2 \dots C_{m+n}) \quad (4)$$

其中,

$$P_r(C_i) = \begin{cases} ps_i, i = 1, 2, \dots, 6 \\ psw_{i-6}, i = 7, 8, \dots, 14 \end{cases}, P_r(C_i C_j) = \begin{cases} ps_i \cdot ps_j, i, j = 1, 2, \dots, 6 \\ ps_i \cdot psw_{j-6}, i = 1, 2, \dots, 6, j = 7, 8, \dots, 14, \dots \\ ps_{j-6} \cdot psw_{j-6}, i, j = 7, 8, \dots, 14 \end{cases} \quad (5)$$

进一步求得, 机房网络连通的可靠度为:

$$P(S_1, SW_8) = 1 - Q(S_1, SW_8) \quad (6)$$

利用 Python 程序进行问题求解, 得到最终网络可靠度为: $P(S_1, SW_8) = 0.1304$, 该结果也验证了前面的初步猜测, 这是因为在该实验机房网络中, 所有设备之间的连接均存在失效的可能性, 整个网络的数据传输在每一条链路上都有可能丢失数据信息, 这使得网络中两终端之间通信是极其不可靠的。在这个网络结构中, 我们可以得到虽然网络中每一个设备失效的概率比较低, 但是综合起来考虑仍然会对整个群网络的通信造成重大影响, 得到较低的传输可靠度。

4 计算机模拟验证实验

4.1 探究网络可靠度与设备失效概率分布之间的关系

通过前文理论分析与实例考证, 可以初步发现, 网络系统中其传输可靠性与网络中各设备的失效概率存在一定的关系。现利用计算机进行模拟仿真实验, 构建一个虚拟的实验网络, 按照指定概率分布随机生成所有设备的失效概率, 进行 100 次模拟实验。对每一次实验均计算其网络可靠度, 最终研究网络可靠度与设备失效概率分布时间的关系, 具体实验结果见图 1:

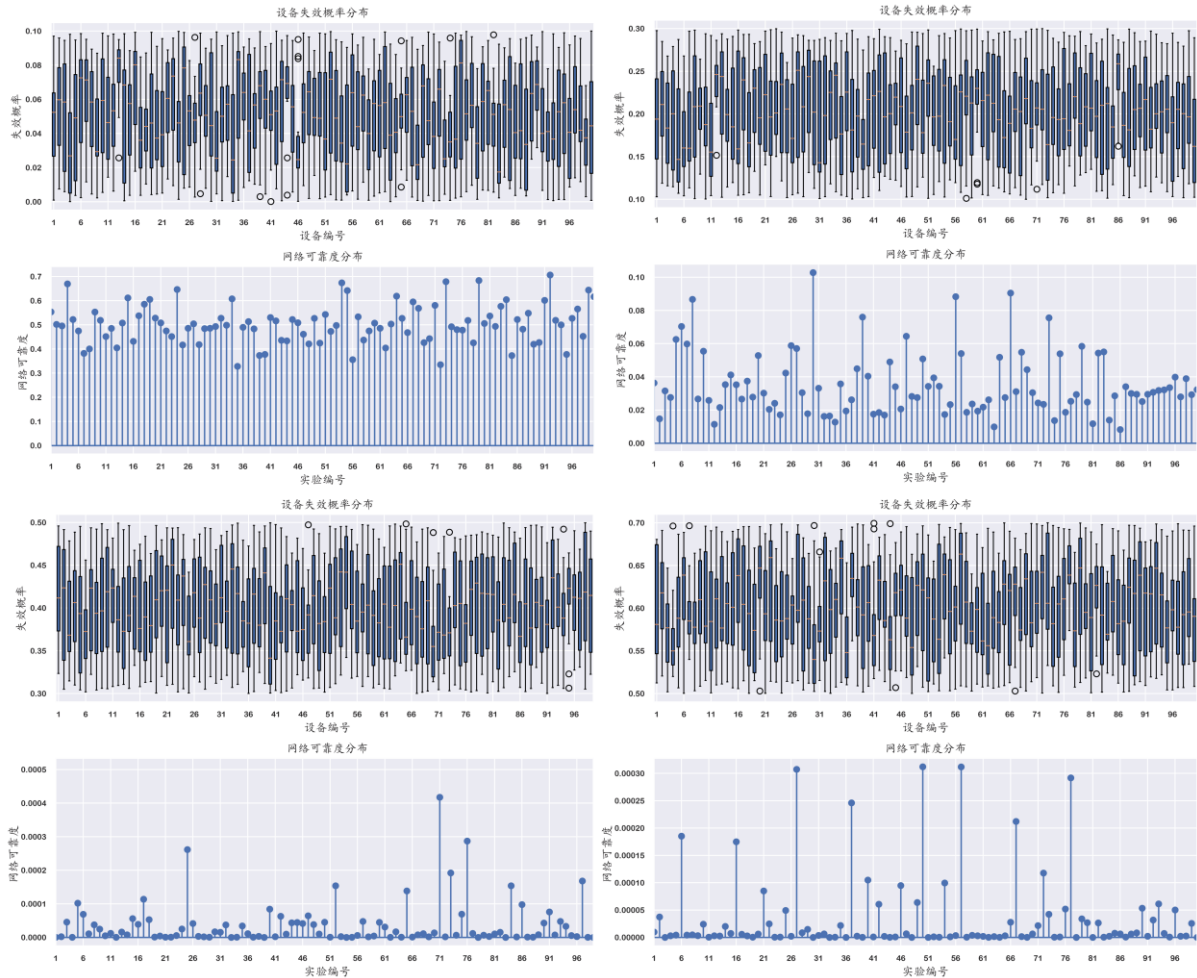


图 1 网络可靠度与设备失效概率分布之间的关系探究

以上四次实验结果各自的设备失效概率分布为:

表 3 网络设备失效概率分布

实验 批次	Exp1	Exp2	Exp3	Exp4
概率 分布	$P_r(C_i) \sim U(0,0.1)$	$P_r(C_i) \sim U(0.1,0.3)$	$P_r(C_i) \sim U(0.3,0.5)$	$P_r(C_i) \sim U(0.5,0.7)$

从以上结果分析可以得到,在网络系统中设备的失效概率越高,所引起的网络可靠度将会越低,并且这种变化的趋势非常迅速,也从中看出对于一个互联型网络,如果存在若干设备有较高的失效概率,很可能引起整个网络系统的瘫痪,也表明网络系统对部分关键设备存在较强的依赖性。

4.2 探究网络可靠度与链路失效概率分布之间的关系

在网络系统中,可靠的网络传输不仅与网络设备的正常运行密切相关,也与设备之间的链路是否能够正常传输数据有着紧密联系。因此,本文继续探究网络可靠度与链路失效概率分布之间的关系,通过计算机仿真实验,随机模拟 100 次实验,并且假设网络系统结构是全连接的即设备之间均存在链路连接,具体实验结果见图 2:

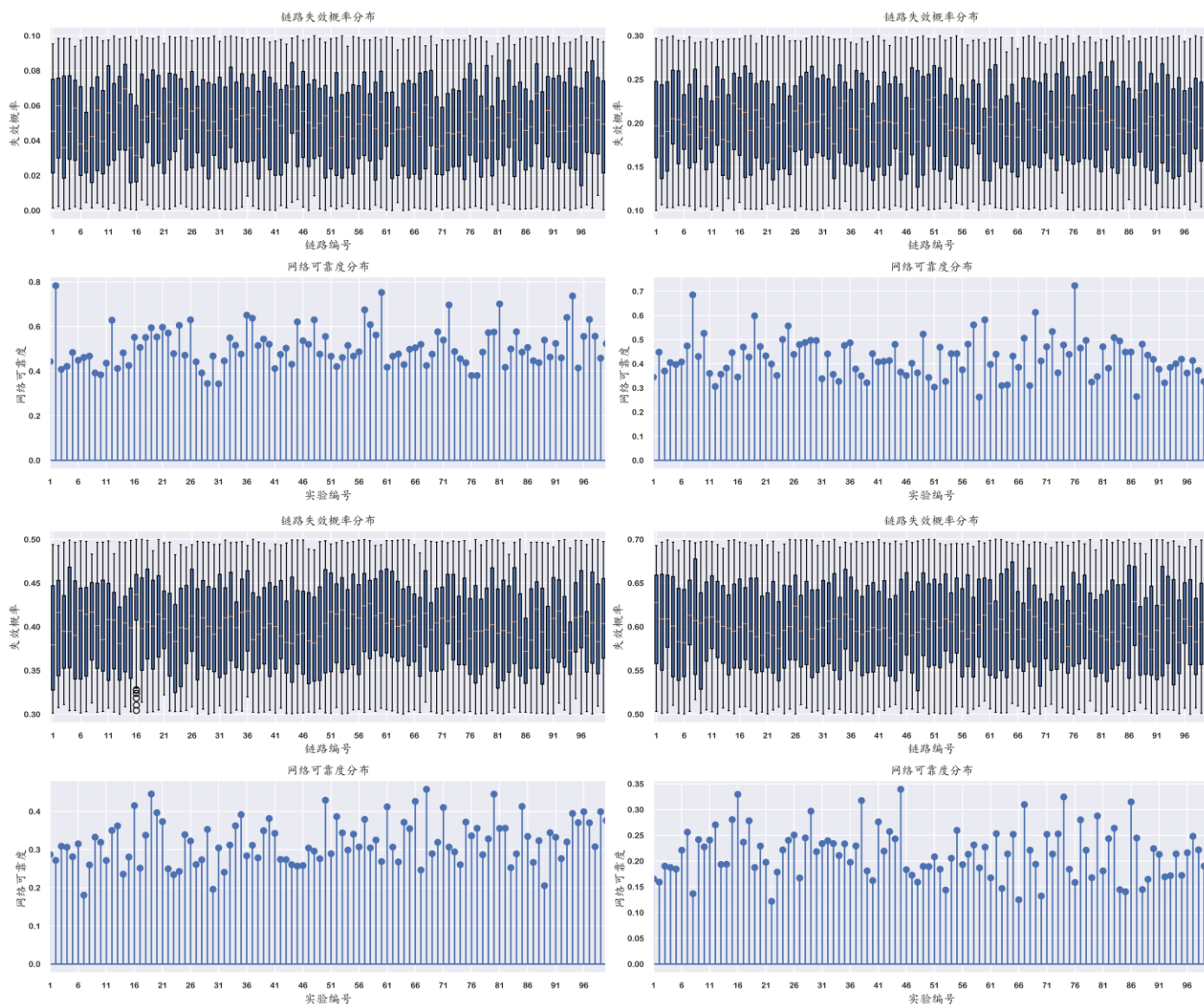


图 2 网络可靠度与链路失效概率分布之间的关系探究

以上四次实验结果各自的设备失效概率分布为:

表 4 网络链路失效概率分布

实验 批次	Exp1	Exp2	Exp3	Exp4
概率 分布	$P_{link} \sim U(0,0.1)$	$P_{link} \sim U(0.1,0.3)$	$P_{link} \sim U(0.3,0.5)$	$P_{link} \sim U(0.5,0.7)$

从以上结果分析可以得到，在网络系统中链路的失效概率越高，所引起的网络可靠度将会越低，但是影响程度比所有网络设备对网络可靠度的小，一种可能的原因是在网络设备均存在失效的概率时，网络是无法与之进行通信的；然而在全连接网络结构中，某些链路如果失效，但是存在到设备的可达链路，那么网络系统仍能够与之进行通信。这种特性，会使得网络链路失效对网络系统可靠度的影响弹性相对网络设备而言较小。

5 结 论

本文通过组合数学中容斥原理，利用传统的解析算法探究了通信网络在数据传输中可靠性的评估，对网络的可靠度进行了分析与量化，并提出了一种高效的求解算法。在实际应用场景中，本文以高校实验机房的计算机网络系统为例进行着重分析计算，结合实际情况验证了理论的正确性；通过实例初步发现网络可靠度与网络设备失效概率分布和网络链路失效概率分布存在紧密联系，因此本文通过计算机仿真实验，模拟探究了网络可靠度与网络设备失效概率分布和网络链路失效概率分布之间的联系，得到了如下结论：

1. 在网络系统中设备的失效概率越高，所引起的网络可靠度将会越低，并且这种变化的趋势非常迅速；如果存在若干设备有较高的失效概率，很可能会引起整个网络系统的瘫痪；
2. 在网络系统中链路的失效概率越高，所引起的网络可靠度将会越低，但是其影响程度比所有网络设备失效对网络可靠度的影响要小；
3. 在全连接网络系统中存在功能互补的链路，能够弥补失效链路的功能，使得网络系统仍能够与某些设备进行通信。

参考文献:

- [1] 李瑞莹, 康锐. 网络可靠性评价研究综述[J]. 可靠性工程, 2008, 7(3): 104—108, 136
Li Ruiying, Kang Rui. A review of research on network reliability evaluation[J]. Reliability Engineering, 2008, 7(3): 104-108, 136
- [2] 孙艳蕊, 赵连昌, 张祥德. 计算网络可靠度的容斥原理算法[J]. 小型微型计算机系统, 2007(05): 830-833.
Sun Yanrui, Zhao Lianchang, Zhang Xiangde. Inclusion-exclusion principle algorithm for calculating network reliability [J]. Small Microcomputer Systems, 2007(05): 830-833.
- [3] 孙艳蕊, 张祥德. 一个计算无圈有向网络可靠度的有效算法[J]. 东北大学学报, 2004(11): 1038-1041.
Sun Yanrui, Zhang Xiangde. An effective algorithm for calculating the reliability of acyclic directed networks [J]. Journal of Northeastern University, 2004(11): 1038-1041.