文章编号：2096-1472(2016)-12-01-04

计算机网络中的故障定位研究

周 畅
（华东师范大学通信工程系，上海 200241）

摘 要：随着信息技术的快速发展，计算机网络日趋复杂，故障定位技术作为网络管理的核心一直是研究的热点。由于网络的互联性，网络故障存在着一定的关联，而计算机网络故障定位则是需要依据事件的相互关系，从多个故障事件中定位故障源。提出专家系统技术，基于主动轮询技术，基于拓扑结构，图论技术对此课题展开研究，并且详细介绍了基于蚁群的高效优化算法，大大提高了工作效率，为之后的故障定位技术的研究有着重要的借鉴意义。

关键词：故障定位；专家系统技术；网络拓扑结构；图论；蚁群算法
中图分类号：TP393.0 **文献标识码**：A

A Study on the Fault Localization Technology in the Computer Network

ZHOU Chang
(*East China Normal University*,*Shanghai* 200241,*China*)

Abstract: With the rapid development of information technology,the computer network becomes more complex.The fault localization technology,as the core of network management,has been a hot research.Because of the connectivity of the network,there are some correlations among network faults.Based on the correlations among events,the source should be localized from multiple computer network faults.The study is conducted based on the expert system technology,the active polling technology,the network topology and the graph theory.The paper proposes a highly efficient optimization algorithm based on ACA(Ant Colony Algorithm),which greatly improves work efficiency and has great reference significance for future studies on fault location technology.

Keywords:fault localization;the expert system technology;the network topology;the graph theory;ACA

1 引言(Introduction)

随着各路网络应用的兴起，用户对于服务质量有着更高的要求。最关键问题之一就是维护计算机网络的可靠性。在计算机网络中故障是不可避免的，因此立即定位与处理是十分关键的[1]。故障管理过程通常分为四个阶段：故障检测、故障定位、修理和测试。其中的故障定位是为了解释报警而分析的一系列故障暗示，准确地发现故障的位置[2]。为了提高网络的可靠性，能够迅速，准确地定位故障尤为重要。尽管前人对故障定位技术进行了不懈的研究，这些方法是从计算机科学的不同领域衍生出来的，包括人工智能、神经网络、信息论和自动化理论。但目前为止在复杂的通信系统中故障定位仍然是一个难题，还有许多问题有待于解决，本文则通过研究多种分析方法来实现故障定位。

2 专家系统技术(The expert system technology)

在故障定位和诊断中应用最为广泛的就是专家系统技术。专家系统试图去反应人类专家的行为，它们基于模仿一个人的，可能是从经验，或是基于它们的原则去理解系统行为。

2.1 规则库系统

仅仅依赖表面知识的方法是基于规则的推理系统，它不需要去深入理解底层系统结构和操作原则。所以对于小系统而言，它则可以提供一个强有力的工具去消除最不可能的假设。但是规则库系统也有许多缺点，例如难以更新知识，以及难以维持。因为通常系统包括硬编码的网络配置，基于观察得出的统计数据虽然已经自动派生出相关规则，但是当系统配置被改变时一定会再生出大部分的相关规则，可见规则库系统是效率低下的和不能处理不精确的情况。缺乏结构的规则库系统通常很难允许使用在分层构建的分布式系统中。

2.2 案例库系统

案例库系统是专家系统的一个特殊类别。基于经验以及过去情况，它试图通过之前解决方案的相关信息去处理被提出的问题。当一个问题被解决时，解决方案可以被用来解决后续的问题。然而，在解决过程中需要一个应用程序特定的模型，以及效率低可能使我们不能用于报警相关。

2.3 决策树

决策树是通过用户观察到的症状去定位问题的根本原因的方法。它是对于专家知识的一个简单表达性的展现。然

而，它的使用性受限于具体应用，以及因存在噪声使其准确度退化。

3 主动轮询技术(Active polling technology)

主动轮询技术是网络管理过程中主动去轮流查询整个计算机网络中设备的各个状态，即去访问简单网络管理协议代理的信息管理库，并且等待系统响应。如果所得的返回结果正常，则只是简单的将查询结果存档。但是如果等到超时，还是没有结果，则是系统发生故障，需要管理员发出告警信息，进行故障定位。

其中的定时轮询用到了VC中MFC类库中的窗口时间事件响应函数，即CSnmpWnd::OnTimer(UINT nIDEvent)。在窗口的初始化函数BOOL OnInitDialog()中添加实现函数SetTimer(nIDEvent,time,NULL)来设定轮流查询的周期，接着在事件函数中添加轮流查询的具体代码，如下：

```
voidCSnmpWnd::OnTimer(UINT nIDEvent)
{CWnd::OnTimer(nIDEvent);
m_GetSnmpHistoryList.AddString(strAsyncGetRequest(m_strOid));}
```

主动轮询技术具有全面，可靠收集网络信息的特点，可以发现网络故障，进行诊断和定位。但是虽然所掌握的网络参数以及状态全面，但是耗时长。在准备阶段必须权衡故障查询速度与所占网络带宽。显然，故障查询速度越快，网络带宽消耗越大，将会直接影响到通信系统地正常运行。

4 网络拓扑结构(Topological structure)

为了实现故障定位，我们将故障定位系统的体系结构分为四个模块，分别为拓扑发现、告警收集、告警关联及故障定位模块。

拓扑发现模块的目的是用于自动发现单个管理域，以及Internet主干网络拓扑，并且尽可能减少对网络的假设条件。按发现拓扑图的方法可分为三种。一是基于探测程序，通过开发专门的探针程序分布在计算机网络中，在程序之间相互测量，从而获得一些关于网络拓扑的信息。二是利用一些通用协议构建拓扑图。三是利用网络设备保存的实际数据，其中利用SNMP协议获得网络设备中保存的MIB数据构造拓扑图是非常有效的拓扑发现方法。

当一个网络设备发生故障时，可产生大量的告警信息。这些告警信息可以用于故障定位。可是当告警过多致使网络管理中心瘫痪时，就需要使用某种机制减少告警信息，因此告警相关的出现解决了这个问题。告警关联则是通过告警信息在时间与空间上进行相关处理，从而减少告警信息数，大大减少了网络故障修复的时间。告警关联从时间与空间两个方面展开。时间方面是针对每一个告警携带的时间戳，从时间序列角度关联各个告警序列。而空间方面则是从网络拓扑结构中关联告警序列。因此，可以通过依赖搜索树模型实现告警关联。其主要思想是根据各节点的依赖关系，将网络拓扑结构构造成树状结构，其中树状结构中的每一个节点都记录着与其直接依赖节点的相关信息。而由于计算机网络拓扑的复杂性，往往构造出的树状是森林。但对于特定的计算机网络而言，树状结构变化非常小，因而我们可以将此作为搜索节点信息的依据，从而得到其中隐藏拓扑结构中的信息。

接着运用贪心算法来处理告警信息，实现故障定位。贪心算法的核心思想则是通过问题的局部最优解来解决整个问题的最优解，即先将一组数据排序找出最优值，进行分析处理，再找出最优值，进行分析处理，直到得到最期望的结果，找到告警信息的最佳解释，从而实现故障定位的目的。

根据目前的网络管理协议(如SNMP)，每个征兆被独立报告，并且由于一个故障源可能导致大量故障征兆，所以很难在大量征兆中找到故障源。我们发现大量的告警信息之间存在相关性，于是对告警信息进行关联处理之后，再进行故障定位。此方法为计算机网络提供可靠性高，实用性强的网管平台，对于故障定位方面具有重要的意义。

5 图论技术(Gragh theory)

图形理论技术依赖于一个系统图形模型，该模型被称为故障传播模型。该故障传播模型包括所有故障及发生在系统中故障症状的全部体现，其中的节点表示故障症状。故障传播模型采用因果图与依赖图的形式。

因果关系图是一个有向无环图$G_c(E,C)$，其节点E表示事件及边缘C表示事件之间的因果关系。边$(e_i,e_j)\in C$表示事件e_i引起事件e_j的事实，被表示为$e_i\rightarrow e_j$。因果图中的边可以通过因果的概率被标记。而依赖图是有向图$G=(O,D)$，其中O是有限的非空集合的对象，D是对象之间的一组边。有向边$(o_i,o_j)\in D$表示o_i中的错误或故障可能导致o_j中的错误的事实，同样可以在边上设置概率值来表现其依赖强度。

5.1 上下文无关法

上下文无关法是允许从子表达式中构建表达式，即从已定义的对象中构建出复合网络对象，从而用于分层组织通信系统。对于有限类的上下无关模型的定位故障源问题在语义上等价于依赖图模型，可以被转换为0—1整数线性规划问题。

5.2 码本技术

码本技术的故障传播模型是由问题码的矩阵表示。在确定性技术中，码字是由0、1序列组成。第i个位置的码字1表示问题p_j与症状s_i的因果关系。码本技术中认为数据是在一个离

散无记忆有损信道上进行传播，其输入的字母表是一组最优化代码以及输出的字母表是一组所有可能症状的集合。有了这样的解释，相关事件等效于将一个已经收到的输出符号解码成有效的输入符号。在相关事件问题中已经收到的符号也来自于{0,1}序列，其中1表示出现特定症状，而0则表示没有观察到特定症状。信道错误导致症状丢失或是虚假，码本以及解码方案决定了可被检测和纠正的错误数量。

码本技术运用最小符号距离作为决策方案。在概率模型中，$d(a,b)$表示两个概率$a,b\in[0,1]$之间的距离，被计算为$\log(a/b)$。其中$\log(0/0)=0$和$\log(a/0)=a$。一旦执行编码，码本技术是非常有效的。然而，它的计算复杂度被$(k+1)\log(p)$所限制，其中k是解码阶段被矫正的错误数目，p是故障的数量。当有多个故障产生，或是发生重叠时，码本技术的精确度很难被预测。此外，由于配置系统的改变需要再生码本，可见这是一个耗时的过程，因此，码本技术不适用于动态改变依赖关系的环境。

5.3　贝叶斯置信网模型

基于传感器数据以及人们的感知信息的贝叶斯网络是固定的，每个贝叶斯网络包括两层：故障层和故障症状层[3]。贝叶斯在2015年提出了一种基于信号观察与在正常情况下信号重建之间差异的故障诊断的研究[4]。并且贝叶斯置信网对于故障定位也有重要意义。贝叶斯置信网实际上是基于概率的不确定性推理网络，其网络拓扑结构就是用概率来表示变量之间相互关系的有向无环图。其中节点表示变量，有向边则表示变量之间的相互关系。贝叶斯也提出：在推理过程中，知识并不是以联合概率分布形式表现的，而是以变量之间的相关性和条件相关性表现的，即用条件概率表示。在故障定位中，根据Pearl理论，对节点$x=\{x_1,\cdots,x_i,\cdots,x_n\}$的概率$P(x)$可以表示为$x_i$与其父节点之间一系列条件概率与边缘分布的乘积，即

$$P(x)=\prod_{i=1}^{n}P(x_i|Parents(x_i))$$

贝叶斯网络是一种网络变量关系的直接表示，而不是推理过程。因此，有向边所指的方向是真正的关系方向，而不是推理过程的流向。

6　蚁群高效优化算法(Highly efficient optimization algorithm based on ant colony)

在这一节中，针对计算机网络故障定位，一种基于蚁群的高效优化算法被提出。最近，在解决工程问题中，群集智能的方法引起了广泛的关注，基于连接的蚁群优化也已经成功应用到工程领域的研究中。蚁群优化是一个多代理系统，他有许多功能。例如分布式长期记忆的使用，强化学习模式的相似功能，以及基于随机组件的一种整体与局部的搜索能力。

我们的算法不同于现有的网络故障研究定位技术，因为它结合了主动与被动测量，通过发送与终止数据来检测故障，以及用主动测量来定位故障。蚁群优化相对于其他的主动测量都要好，因为蚂蚁可以智能行动，因此可以更加高效的定位故障节点[1]。

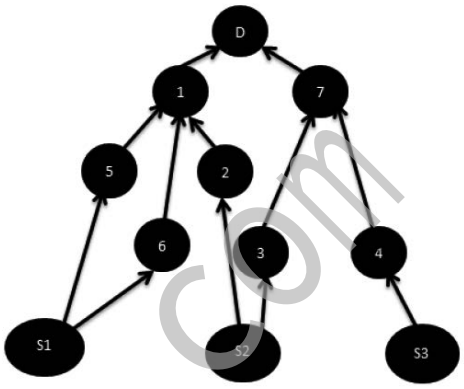

图1 至少一个节点的网络拓扑实例

Fig.1 An example of a network topology that it has at least a faulty node

考虑一个网络，数据从客户端发送到服务器，于是我们考虑到端的数据就可检测网络中的故障组件[5]。在所提到的算法中主要有两个方面。第一，一组节点作为目标节点(服务器)和一组节点作为源(客户端)，然后基于主动测量选择一组节点移动到目标节点。在目标节点中通过蚂蚁估计每个节点的故障。第二，候选节点通过被动测量被用于测试。下面将对每一步做详细的说明。

将一个或多个服务器视为目标节点，以及客户端被认为源节点，源节点被认为是蚂蚁的巢，蚂蚁需要移动到服务器。每个蚂蚁由二进制数组组成，数组的长度是所有网络节点的数量。数组的每一个元素表示网络中的节点，这些元素仅由0、1两个数值表示。当一个蚂蚁从源节点移动到目标节点，过程中通过的每一个节点的值在数组中都等于1。在图1中，有一个目标节点($D=1$)和三个源节点($S=3$)，以及七个在它们之间的节点($N=7$)。一共有五只蚂蚁，分别为a_1、a_2、a_3、a_4、a_5。可以看出蚂蚁的数量由源节点上输出边的数量决定。当蚂蚁通过一个节点时，蚂蚁数组就被刷新。当蚂蚁到达目标节点时将有五个数组，为$\{A_1,\cdots,A_5\}$。用$S_1=\{a_1,a_2\}$,$S_2=\{a_3,a_4\}$和$S_3=\{a_5\}$分别表示每个蚂蚁属于的一个源(S_1、S_2或S_3)，从一个公共源出发的蚂蚁进行OR逻辑运算符操作。因此，在上面的例子中将有三个数组。图2展现的是OR操作符之后的数组，即每个数组表示同一个源的一系列节点。

表1 被存储目的数组的数值

Tab.1 Values stored in the destination arrays

数组名称	数组内容					
S_1	0	0	0	1	1	0
S_2	1	1	0	0	0	1
S_3	0	0	1	0	0	1

根据以上描述，对于每个节点的故障概率为p_i，通过方程式(1)得到。其中N是出现的网络节点的总数。对于每个节点可以通过方程式(2)获得m_i，m_i是数组的节点。在等式(2)中$S(j,i)$表示第j个数组的第j个元素。例如，在等式(2)中，m_1=2，m_2=1，m_3=1，m_4=1，m_5=1，m_6=1和m_7=2。通过等式(1)可以计算出P_1=0.22，P_2=0.11，P_3=0.11，P_4=0.11，P_5=0.11，P_6=0.11和P_7=0.22。

$$P_i=m_i/(\sum_{i=1}^{N} m_i) \tag{1}$$

$$m_i=\sum_{j=1}^{N} S(j,i) \tag{2}$$

在计算出故障概率后，应用蚁群优化选择最佳节点用于测试。蚂蚁开始从客户端(源节点)到服务器(目标节点)，一个基于故障概率为p_i及节点上具有信息素的蚂蚁需要选择将要移动到的下一个节点。由于故障概率与信息素互相依赖，所以蚂蚁可以依照以下两个标准去选择接下来的节点：(1)第i个节点的故障概率；(2)在下一个节点的信息素浓度。首先，假设所有节点的信息素浓度都相同，于是可以基于等式(3)，蚂蚁选择具有最高p的下一步去移动。

$$S=\begin{cases} S_1=\arg(\max[(phermone(n_i))*(p_i)^{\beta}]) & if\ q \le q_0 \\ S_2=p \quad otherwise \end{cases} \tag{3}$$

$$P_k=(phermone(n_i)*p_i\ \wedge\beta)/(\sum_{i=1}^{N} phermone(n_i)*p_i\ \wedge\beta) \tag{4}$$

信息素(n_i)表示第i个节点上信息素的浓度。B是衡量信息素与p_i相关性的参数。q是以[0,1]均匀分布下被随机选择的值。q_0($0\leqslant q_0\leqslant 1$)是一个参数，是针对有高概率$p_i$与高信息素的节点。$S$是蚂蚁$k$选择移动到下一节点的概率。

一般来说，信息素被全局或局部更新所改变，即当蚂蚁从一个节点转移到另一节点时，信息素依据等式(4)所更新。

$$phermone(n_i)=(1-\propto)\cdot phermone(n_i)+\propto\cdot p_i \tag{5}$$

在等式(5)中，$\propto$是0与1之间的数。如果$\propto$=1，则信息素更新取决于故障概率。如果$\propto$=0，那么信息素的更新取决于在n_i信息素的浓度。$\propto$在0与1之间取值，该值决定了故障概率与n_i信息素浓度对于信息素的影响强度。

蚁群算法结束后，具有更多信息素的节点被选为测试。确定蚁群算法的最后条件是在每次巡回(蚂蚁从客户端到服务器叫一次巡回)的最后，信息素的浓度是否有影响。如果节点上的信息素在几次连续重复中没有改变，则应该被停止，从而选择一个具有更高信息素浓度的节点。本文所提出的蚁群算法可以在迭代中运行直到所有的故障组件被定位。图3展现了蚁群算法的流程图。

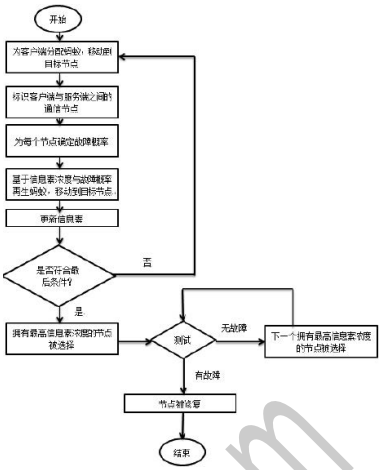

图2 所提出的算法的流程图

Fig.2 Flowchart of proposed algorithm

7 结论(Conclusion)

随着计算机网络应用的大规模普及，网络已经成为我们生活中不可或缺的设备。与此同时，网络的复杂性也日益加大，因此检测定位并排除故障，保障用户网络安全与通畅尤为重要。本文提出了基于专家系统技术、主动轮询、网络拓扑、图论的故障定位方法，并且着重讨论了一种基于蚁群的高效优化算法。利用蚁群算法良好的正反馈与容错性的特点，对复杂的计算机网络进行故障定位。该算法大大减少了操作者的运算量，并且使网络管理速率大幅度提高。

参考文献(References)

[1] GARSHASBI M S.Fault Localization Based on Combines Active and Passive Measurements in Computer Networks by Ant Colony Optimization[J].Reliability Engineering and System Safety,2016,152:205-212.

[2] Bing W,et al.Fault Localization Using Passiveend-to-End Measurement Sand Sequential Testing for Wireless Sensor Networks[J].IEEE Trans Mobile Comput,2012,11:439-452.

[3] CaiB,et al.Multi-Source Information Fusion Based Fault Diagnosis of Ground-Source Heat Pump Using Bayesian Network[J].Apply Energy,2014,114:1-9.

[4] BaraldiP,et al.Comparison of Data-Driven Reconstruction Methods for Fault Detection[J].Reliab,IEEE Trans,2015,64(3):852-860.

[5] SteinderM,AdarshpalS.Probabilistic Fault Localization in Communication Systems Using Belief Networks[J].IEEE/ACM Trans Netw,2004,12:809-821.

作者简介：

周 畅(1996-)，女，本科生.研究领域：通信工程.