

下一代网络协议测试数据半自动生成方法研究^{*}

李 华 叶新铭 吴承勇 王 龙
(内蒙古大学计算机学院 呼和浩特 010021)

摘 要 在协议测试过程中,对于测试数据的生成方法的考虑应从包的粒度细化到字段上。对协议的通用数据包进行了全面的分析,根据各个字段的功能及其取值之间的关系,对所有字段进行了固定字段、独立字段以及相关字段序列的划分,然后根据测试经验对独立字段和相关字段序列进行了优先级赋值,并根据优先级构建测试包,从而可以将具有重要测试目的的包尽早生成进行测试,并根据优先级标记指导测试结果判定。给出了测试包库的生成算法,并分析了该算法的有效性,然后以 BGP4+ 的 UPDATE 包为例给出了生成的包,最后给出了结论。

关键词 协议,一致性测试,包,字段

Research on Semi-automatic Generation Method of Testing Data for Next Generation Internet Protocol

LI Hua YE Xin-ming WU Cheng-yong WANG Long
(School of Computer Science, Neimongol University, Hohhot 010021, China)

Abstract During the process of protocol testing, the consideration of test data should be changed from the granularity of packet to the granularity of field. The general format of packet was analyzed completely. According to their functions and the relation of the values, the fields were divided into three parts, the first part is constant part, the second part is independent fields and the third part is the related fields. Moreover based on the accumulated experience, the priority was assigned to the independent fields and each sequence of related field as a whole. The testing packets were sequentially constructed according to priorities such that the important packets are generated and some aspects of the protocol are tested as early as possible. The tag of priority is used to conduct the verdict of testing result. The generation algorithm of packet database was provided and its validation was analyzed. The UPDATE of BGP4+ was presented as an example to verify the algorithm. Finally the conclusion was presented.

Keywords Protocol, Conformance testing, Packet, Field

1 引言

随着下一代因特网的部署,与之相关的协议开发工作进展迅速,因此许多基本的测试工作急需实施,其中就包括下一代因特网相关协议的一致性测试。通过对下一代网络相关协议分析可知,虽然所有的与 IPv6^[1] 相关的协议都与前一版本协议具有相同的有限状态机模型,但相应的包的格式却变化很大。文献[2]分析了具有相同形式模型的协议测试所面临的问题以及测试方法,给出了 OSPFv3^[3] 相比 OSPFv2 的主要变化,并以 OSPFv3 的 Hello 包为例阐述了具体的测试方法。在实际测试过程中,如何构建有意义的和高效的测试数据包,以便在协议测试中较早发现测试实现的问题,继而加快协议的开发速度以及路由器投放市场的速度,就成为协议测试研究的重点。本文的组织结构如下,第 2 部分介绍常用的测试方法,并简介现有的一些测试工作,第 3 部分介绍协议测试包生成方法及其效率分析,第 4 部分以 BGP4+ 中的 UPDATE 包为例验证给出的方法,最后是本文的结论。

2 协议的一致性测试方法分析

验证协议实现是否与其标准相一致的过程称为一致性测试,它是保证协议实现质量的一个必需而且十分重要的手段。ISO 制定了协议一致性测试标准 ISO/IEC-9646^[4],这个标准是针对使用自然语言描述的协议测试而制定的,它根据不同的控制观察点定义了不同的抽象测试方法。现有的端系统抽象测试方法基本上可以分为 4 类,即本地测试法、协调式测试法、分布式测试法和远程测试法。由于远程测试法的主要特点是不要求能访问 IUT(被测实现)的上边界,也不要求显式

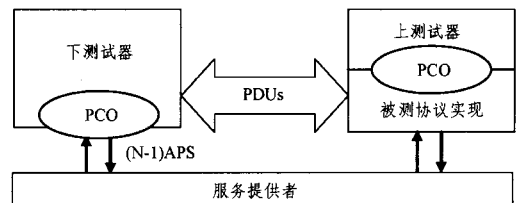


图 1 远程测试法

^{*} 国家自然科学基金项目(60263002),内蒙古自然科学基金重点项目(200711020803),中欧联合项目 GO4IT。李 华 教授,硕士生导师,CCF 会员,研究方向为分布式系统与网络计算;叶新铭 教授,博士生导师,主要从事计算机网络和分布式系统方面的研究;吴承勇 教授,硕士生导师,主要研究方向为计算机网络与电子商务;王 龙 硕士,主要研究方向为协议测试。

地测试协调过程,只是把 IUT 当作黑盒进行测试。测试判决是由基于下测试器对 IUT 提供的激励以及下测试器所观察到的 IUT 的响应做出的,所以我们在协议测试研究中,选择了远程测试方法,如图 1 所示。

目前关于测试生成方法的研究多数是集中于有限状态机、带标记转换系统等基于形式描述方法的研究,如文献[5,6]基于有限状态机,提出了产生测试序列的方法,然后进行测试框架设计,根据测试目的进行测试配置,最后选择测试数据,执行测试用例,得到判定结果。就网络协议测试而言,特别是对于与下一代网络相关的协议,它们对应的有限状态机模型与支持 IPv4 时并没有变化,但各类包格式的扩展却依据不同的协议而不同,如 BGP4+^[7]与 BGP4 相比,其变化主要是扩展或者改变了 UPDATE 包中某些字段的功能。在文献[8]中,采用了一种基于保证错误覆盖的测试用例生成方法,首先根据概率算法选择可能的网络配置,然后在其上做测试,测试中用的网络拓扑结构是依据概率方法生成的。该文检查了网络拓扑数据库、路由表以及包转发行为,提出了基于概率的测试方法,同时给出了路由协议的测试策略,并对 3 个路由协议(如 RIP, OSPF, BGP)给出了测试算法,最后给出了一个测试框架。文献[9]研究了邻居发现协议的测试方法,给出了邻居发现协议的 MSC 模型,基于 TTCN2 进行了一致性测试研究,文[10]分析了 BGP4+ 协议,并对它的互操作性测试进行了较为全面的研究并进行了相关的实验。在实验过程中,我们发现了测试数据选择的重要之处,因此本文对测试数据的生成进行了较详细的研究。

对于 IPv6 以及相关协议而言,大部分的扩展都是在 IPv4 相应协议的基础上进行的。在扩展之后的协议中,由于协议说明中的有限状态机与前一版本的有限状态机一样(如果该协议在说明中描述了有限状态机),对于扩展部分的测试及其自动化问题就成为急需解决的问题。而在这个问题中,如何快速生成有意义的测试包则是测试过程中一个关键的步骤。本文就是针对这一问题,以 BGP4+ 为例,说明了自己的测试方法。

3 协议测试数据生成问题分析

对于协议测试而言,针对不同的测试目的,测试方需要组建不同的数据包。如何自动生成有意义的数据包,对于被测进行有一定针对性的测试。

3.1 通用数据包字段的分析

在网络的工作过程中,所有的信息交互都是以数据包的方式进行的。一般一个数据包由多个字段构成,并且依据测试目的的不同,对于测试结果所关注的内容也会不同。也就是说,在测试过程中,可以将包内字段按所关注的优先级进行划分。采用这种方法可以避免盲目构建各类包,从而提高测试效率,降低测试费用。通过对以往测试过程以及测试目的的分析,可以得出,测试目的的最终体现是由生成的测试包的某个或某几个字段实现的,所以应首先对包的格式进行分析。

在任何协议中,数据包的一般格式都如图 2 所示。

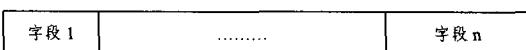


图 2 包的通用格式

假设用一个多元组 $p=(p_1, p_2, p_3, \dots, p_n)$ 表示一个包, p_i 表示包中某个字段, $|p|=n, 1 \leq i, k, l \leq n$, 且 $i \neq k \neq l$, 其元素的关系是:

(1) p_i 的值是固定的,如数据包中的版本号等。

(2) 相关的。如果 p 中的字段 p_i, p_k 是相关的,当且仅当 p_i, p_k 的取值相关,即当 p_i 取某值时, p_k 的取值也随之确定,并且多元组中无 p_i 时,也无 p_k ,此时称 p_i 与 p_k 为 2 度相关的;当 p_i 取某值时, p_k 和 p_l 的取值也随之确定,并且多元组中无 p_i 时,也无 p_k, p_l ,称 p_i 与 p_k, p_l 为 3 度相关的, p_i 称为相关字段的关键字段。当然还可以根据需要定义更高的相关度。相关关系具有不可逆的特点,即如果 p_i 和 p_k, p_l 是 3 度相关的,并不意味着 p_k 和 p_i, p_l 一定是 3 度相关的,因为此处 p_i 是关键字段。

(3) 独立的。如果 p 中的字段是独立的,当且仅当 p_i, p_j 的取值不是相关的。

假设 $p=(p_1, p_2, p_3, \dots, p_n)$ 表示一类包的通用格式,其中 n 表示包中字段的最大数目, p_i 表示字段,并且各不相同,其中 $n \geq i \geq 1$ 。根据测试经验可以对 p 划分为 3 个子字段序列 c, q 和 r ,其中 c 表示固定值序列,其值为常量,由于其值固定,因此在此不多叙述; $q=(q_1, q_2, q_3, \dots, q_m)$, $r=(r_1, r_2, r_3, \dots, r_k)$ 。 q 的各个字段是独立的, r 的字段之间可能有多个相关的关系,如 r_1 和 r_2 是 2 度相关的, r_3, r_4 和 r_5 是 3 度相关的等等。现在对 q 的各个字段和 r 内的各相关字段序列的关键字段进行优先级赋值。优先级对于一个包而言具有局部性,与其它包的格式以及优先级取值无关。优先级值最高为 0,最低为 $n-1$,没有优先级的字段其优先权也是 $n-1$ 。在元组 $(q_1, q_2, \dots, q_i, \dots, q_m)$ 中, q_i 表示构建一个包时需要特别考虑的字段,其值也许不取默认值而需要重新考虑。对于不加下划线的字段,各组之间的同名字段的取值可以相同也可以不同。为了有利于组建通用的包格式,一般假设这些字段的取值相同。为表示方便,在 $(q_1, q_2, q_3, \dots, q_m)$ 中字段的顺序与实际包中对应字段的顺序可能不一致,在构建实际数据包时再加以对应就可以了。

假设独立字段有 j 个元组有优先级, $m-j$ 个元组无优先级,构建集合 Q^* ,其中的每个元素为一个 m 元组,每个 m 元组中只有一个独立字段(即带下划线部分)需要特别考虑。

$$Q^* = \left\{ \begin{matrix} (\underline{q_1} & q_2 & \dots & q_m) \\ (q_1 & \underline{q_2} & \dots & q_m) \\ \vdots \\ (q_1 & q_2 & \dots & \underline{q_{j+1}} \dots q_m) \end{matrix} \right\}, \begin{matrix} |Q^*| = j+1 \\ \text{并且 } j+1 \leq m \end{matrix} \quad (1)$$

为简单起见,假设所有相关度为 2,依据相关性可以将 r 分成 $R^* = \{(r_{i1}, r_{j1}), (r_{i2}, r_{j2}), \dots, (r_{iw}, r_{jw})\}$, 共计 w 个相关组,其中 $w \leq \lfloor k/2 \rfloor$ 。并且如果 R 中还有相关度更大的元组,如相关度为 3 或更高,则 R^* 的元素会更少,即

$$R^* = \{(r_{i1}, r_{j1})(r_{i2}, r_{j2}) \dots (r_{iw}, r_{jw})\}$$

其中: (r_{ix}, r_{jx}) 是相关的, $1 \leq x \leq w$, 并且 $0 \leq$ 每个 (r_{ix}, r_{jx}) 的优先级 $\leq n-1$, r_{ix} 和 r_{jx} 的优先级相同。

$$|R^*| = w \quad (2)$$

然后根据式(1)和(2)重构某类包的集合 P , 得到

$$P = \{P_i = \{Q_s\} \cup \{R_t\} \mid Q_s \in Q^*, 1 \leq s \leq j+1 \text{ 并且 } R_t \in R^*, 1 \leq t \leq w, \text{ 并且每个集合各取一个元素构建 } P_i\}$$

(3)

由式(3)可以得到需要产生的最多包数(为 $(j+1) \times w$),采用这种方法得到的包都是有意义的,也就是说 P 所包含的元组的各字段之间没有矛盾,可用于指导测试包生成。

从定义可以看出,独立字段的取值互不相干,不会产生矛盾;独立字段与相关字段之间也不会有矛盾,二者没有前面定义的相关性,因为如果独立字段与某些相关字段有本文定义的相关性,则该相关字段的相关度可能增 1,从而该独立字段就不应该在 q 的元组中出现。

3.2 包的半自动生成

由于在协议数据包的自动生成过程中,独立字段与相关字段的划分以及优先级是由专家手工参与赋值的,因此我们将自己的方法称为半自动生成方法。为了更容易描述数据包的生成,我们假设有优先级的独立字段以及相关字段的优先级各不相同。

3.2.1 协议数据包的生成算法

首先构建协议包的模板,对每一类包划分独立字段、相关字段。每种包可以分成 3 个部分:固定值部分、独立字段以及相关字段。构建优先级对应表 L_p 和独立字段优先级库 QDB 及其各字段默认值、相关字段优先级库 RDB 及其各字段默认值。

算法 3.1

算法输入:某类包的模板, L_p, QDB, RDB 。

算法输出:带有固定值、独立字段、相关字段的包组成的包集合 PDB 。

数据结构:通用的包格式。

begin

(1) 初始化包模板,填入固定值部分,依据 QDB 以及 RDB 将各个字段的默认值填入;

(2) for Privilege=1 to $j+1$ do
/* j 为优先级
{ 对包中的 q_i 位置的值根据 QDB 赋值,并作标记;

(3) for Related Field=1 to w do
{ (r_i, r_{ji}) = 读入 RDB 相应的赋值;
其它的值保持原默认值;
}

(4) 去掉包中相关字段子序列中各字段值全为 0 的字段,即去掉不应在本次组包中出现的字段

(5) 将输出的包按照优先级顺序放入 PDB 数据库中
}

end

算法分析:(1)部分最大 $O(n)$;(2)部分最大 $O(j+1)$;(3)部分最大 $O(w)$;(4)最大为 $O(n-j)$;(5)部分最大 $O(1)$,所以算法的时间复杂度为

$$O(n) + O((j+1)((w+n-j+1))) \quad (4)$$

由式(1),(2)可知: $j+1+2w=n$;

$$2w+1=n-j \quad (5)$$

则由式(4),(5)可得

$$O(n) + O((j+1)((w+n-j+1))) = O(n) + O(3wj + 3w + 2j + 2) \quad (6)$$

由于 $w+j < n$,由式(6)可得

$$\begin{aligned} O(n) + O(3wj + 3w + 2j + 2) &< O(n) + O(3w(n-w) + 3w + 2(n-w) + 2) \\ &= O(n) + O(3wn - 3w^2 + w + 2n + 2) \\ &< O(n) + O(3wn + 2n) \end{aligned}$$

$$= 3 * O((w+1)n) \quad (7)$$

因为已有 $w \leq \lfloor k/2 \rfloor$,而 $k < n$

$$\text{所以 } w \leq \lfloor n/2 \rfloor \quad (8)$$

由式(7),(8)可得

$$\begin{aligned} 3 * O((w+1)n) &< 3 * O((n/2+1)n) = 3 * \\ &O(n^2/2+n) < 2 * O(n^2) \quad (9) \end{aligned}$$

3.2.2 半自动生成与枚举方法的比较

采用枚举法生成包的方式也可以生成所有的测试数据包,无论这些包是否是合法的测试数据包。假设一个包最多有 n 个字段,每个字段的取值范围是 $v_{i1} - v_{ie}$,并且假设每个字段取值不少于 2,则枚举法生成的测试数据包的数目为

$$\prod_{j=1}^n v_{wj} - v_{ij} \geq 2^n \quad (10)$$

在一般情况下,字段数 n 只要大于 7,根据式(9),(10),就有

$$2 * O(n^2) < 2^n \quad (11)$$

即在字段数超过 7 时,本方法就比枚举方法效率高,并且 n 越大,本方法越好。

4 举例

BGP4+是在 BGP4 的基础上进行了扩展,提供支持下一代网络域间路由协议功能。下面以 BGP4+ 的 UPDATE 包为例介绍各个字段取值,分析它的包结构,介绍上述测试数据包生成方法,然后以一个实际的 UPDATE 包为例,给出它的生成过程。

4.1 UPDATE 包的字段

如果图 3 中 Marker 和 Type 字段的取值分别是全 1 和 2,则表示数据包是 UPDATE 包。UPDATE 包是 BGP 协议中最重要的数据包,与下一代网络相关的内容都在它的包结构中体现。

在 BGP4+ 协议中,Path Attributes (variable) 包括多个路径属性,具体内容根据消息中所需包含的路径属性内容决定。在 BGP4 协议中,有 3 部分与 IPv4 的信息相关,即 NEXT-HOP, NLRI 和 AGGREGATOR 属性。假设任何 BGP 发言者都有一个 IPv4 地址(包含在 AGGREGATOR 属性中),那么要使 BGP4 支持多网络层协议的路由只需要加入两个功能:把特定的网络层协议分别与 NEXT-HOP 和 NLRI 关联起来。为此在 BGP4+ 中增加了两个属性:MP_REACH_NLRI(Multi-protocol Reachable NLRI) 和 MP_UNREACH_NLRI(Multi-protocol Unreachable NLRI)。这两个属性是可选非传递的,不支持多协议扩展的 BGP 发言者会忽略这两个属性中携带的信息并且不转发这些信息。具体 MP_REACH_NLRI 和 MP_UNREACH_NLRI 取值介绍参见文献[7]。

Marker 字段(对于 UPDATE 包,此 16 字节应为全 1)		
Length (2 字节)	Type: 2 (1 字节)	
Unfeasible Routes Length (2 字节)		
Total Path Attribute Length		
Path Attributes (可变量)		

图 3 BGP4+ 中 UPDATE 包的一般格式

4.2 UPDATE 包的半自动生成

假设由主测的测试端口向被测设备发送 UPDATE 包,包中的 Marker 字段(即字段 1)和 BG PType 字段(即字段 3)是取固定值的字段,前者全 1,后者当前字段值为 2(表示是 UPDATE 包)。

表 1 所示的包的测试目的是保证被测设备在接收到外部对等体发来的 Origin 属性为 EGP 的 UPDATE 消息时,能够向其它的对等体转发该路由。相应的独立字段的优先级库 QDB 的相关内容为:

字段 2 Total Length 优先级=6;

字段 4 Unfeasible Routes Length=5;

字段 5 Path Attribute Length=4;

字段 12 Attribute Data Length=3。

上述长度字段(length)之所以列为独立字段是因为它们的值的相关性是可变的。例如字段 5 的值是由 UPDATE 消息具体携带的路径属性来决定的,携带的路径属性不同,字段 5 的值就会发生相应的变化。

UPDATE 包相应的相关字段的 RDB 为 (Attribute Flags, Attribute Type Code, Attribute Data Length, Origin Type),相关度为 4,Origin Type 为该序列的关键字段,优先级设为 0(最高);(Attribute Flags, Attribute Type Code),相关度为 2,Attribute Type Code 为该序列的关键字段,优先级设为 1;(Path Segment Type, Path Segment Length, Path Segment Value),相关度为 3,关键字段为 Path Segment Value,优先级设为 2。

表 1 表示的就是经过算法 3.1 产生的 UPDATE 包的各字段的取值及其含义,也就是发出的 UPDATE 包,图 3 中的可变长路径属性在表 1 中已用大括号括住。本例有 3 个相关字段序列,每个的关键字段已被圈住了(如表 1)。当关键字段的值确定之后,它所在的相关字段序列的其它字段的值随后确定,并且每个相关序列中的各字段具有相同的优先级值。例如字段 9 的值确定之后,字段 6,7,8 的值随后确定,并且与关键字段(即字段 9)的优先级一样,它们的优先级都是 0。其它的相关字段值全为 0 的字段序列已被算法 3.1 去掉。执行测试之后,通过观察侦到的数据包以及查看被转发路由器的路由表,发现被测反应与预测结果相同。

参考 ISO/IEC9646 的建议,根据协议的测试目的和 BGP4+ 协议标准,我们将测试分为(1)互连行为状态机测试;(2)消息格式一致性测试;(3)消息处理功能测试;(4)路由处理功能测试;(5)路由流量控制功能测试;(6)BGP4+ 差错处理功能测试;(7)路由反射功能测试;(8)Community 属性处理功能测试。其中有 84 个测试用例与 UPDATE 有关。根据以往的经验可以知道在编写测试用例的过程中,手工完成数据包填写的工作量是非常大的,并且容易出错。而采用为字段赋优先级的方法构建测试包,不但可以提高组包的速度,还可以保证组包的正确性。例如,UPDATE 数据包的 Marker 字段、Type 字段值可以作为固定值处理。由于 Attribute Flags, Attribute Type Code, Attribute Data Length, Origin Type 字段之间是相关字段,它们的值可能会发生变化,但是如果 Origin Type 的值确定了,则其它 3 个字段的值也就确定了(体现相关性)。在实现了对这些字段的自动、半自动生成后,大大地减少了测试组包的工作量。

表 1 生成的 UPDATE 的各字段取值及其含义

序号	字段名称	值	含义/优先级
1	Marker	16 字节全 1	OPEN 消息中不含有身份鉴别机制,固定值
2	Total Length	92	UPDATE 消息的长度/6
3	Type	2	UPDATE 消息,固定值
4	Unfeasible Routes Length	0	不可行路由长度/5
5	Path Attribute Length	69	UPDATE 消息携带的路径属性的总长度/4
6	Attribute Flags	64	该属性为公认可传递属性/0
7	Attribute Type Code	1	Origin 属性/0
8	Attribute Data Length	1	属性数据占 1 个字节/0
9	Origin Type	①	EGP 表示路由是从外部对等体学习得到的/0
10	Attribute Flags	64	该属性为公认可传递属性/1
11	Attribute Type Code	②	As Path 属性/1
12	Attribute Data Length	6	属性数据占 6 个字节/3
13	Path Segment Type	2	AS_SEQUENCE 路由经过的 AS 的有序集/2
14	Path Segment Length	2	AS_SEQUENCE 路由经过的 AS 个数/2
15	Path Segment Value	64502 64000	路由经过的 AS 号/2

结束语 下一代网络协议与其前一版本的协议在功能上类似,但是在包内容以及结构方面变化较大,通过分析通用包中的字段及其特点,将字段分为独立字段和相关字段,并对独立字段以及各相关字段序列赋予优先级的方法,给出了构建测试包的一般方法,并与枚举方法进行了比较。最后以 BGP4+ 的一个 UPDATE 包为例,给出了通过独立字段和相关字段产生的测试包。本文的方法促进了下一代网络相关协议测试中自动化测试进程,有利于协议测试这一部分工作的完全自动化,减轻测试人员的工作强度,减少出错率,最重要的是引入字段的优先级之后,更容易在测试中较早发现协议实现的致命错误,提高测试效率。下一步的工作将集中在如何让测试的判定部分自动进行,从而提高整个测试过程的自动化程度。

参考文献

- [1] Deering S, Hinden R. Internet Protocol, Version 6 (IPv6) Specification. RFC 2460, December 1998
- [2] 李华,叶新铭,丁雪莲,等. 基于相同形式模型的协议测试方法研究[J]. 内蒙古大学学报:自然科学版,2005,36(6):703-708
- [3] Ferguson C D, Moy J. OSPF for IPv6. RFC 2740, December 1999
- [4] Information technology—Open Systems Interconnection—Conformance testing methodology and framework—Part 1: General concepts. ISO/IEC 9646-1, 1994
- [5] Miller R E, Paul S. Generating Conformance Test Sequences for Combined Control and Data Flow of Communication Protocols// Proceeding of Protocol Specifications[C]. Testing and Verification (PSTV'92). Florida, USA, June 1992
- [6] 王建国,吴建平. 基于扩展有限状态机的协议测试集生成研究[J]. 软件学报,2001,12(8):1197-1204
- [7] Bates T, Rekhter Y, Chandra R. Multiprotocol Extensions for BGP-4 (BGP4+). RFC 2858, June 2000
- [8] Hao Ruibing. Testing IP Routing Protocols from Probabilistic Algorithms to a Software Tool[C]// FORTE/PSTV. Pisa, Italy, 2000:249-264
- [9] 叶新铭,孙美飞. IPv6 邻居发现协议的一致性测试[J]. 计算机科学,2005,32(6):43-46
- [10] 张涛. 边界网关协议 BGP4+ 的互操作性测试研究. 呼和浩特:内蒙古大学,2007