

基于模糊综合决策的服务器推送方法

蒋乾悦 张亚英

(同济大学嵌入式系统与服务计算教育部重点实验室 上海 200092)

摘 要 传统的服务器推送技术不能根据不同的应用场景动态地调整推送方案。提出了将模糊综合决策模型应用于服务器推送的技术,将传统的长轮询技术和轮询技术相结合,提出了一种基于模糊综合决策的复合式轮询技术。并将此推送技术应用于道路交通实时监控平台,使平台在实时信息发布时既能满足不同用户的实时性要求又能充分利用系统服务器资源。

关键词 服务器推送,模糊综合决策,轮询,复合式轮询

中图法分类号 TP393.09 **文献标识码** A

Server Push Technology Based on Fuzzy Fusion Decision-making

JIANG Qian-yue ZHANG Ya-ying

(Key Laboratory of Embedded System and Service Computing of Ministry of Education, Tongji University, Shanghai 200092, China)

Abstract The traditional server push technology cannot change the push method in different user situation. A new server push technology called compound polling was introduced which combines namely polling with long polling based on fuzzy multiple criteria alignment program. Applying this technology to the traffic monitoring platform satisfies multi-user real time requirement as well as full usage of server resource when announcing real time information.

Keywords Server push, Fuzzy comprehensive decision-making, Polling, Compound polling

随着道路上机动车和交通违章事故的增多,交通监管部门和市民都需要及时了解实时的交通信息。交通实时监控平台是一个可以发布交通实时信息的服务平台。传统信息服务平台主要通过两种方式实现信息发布,一种是客户端拽取(Pull),另一种是服务器推送(Push)。客户端拽取是通过客户端用户主动更新页面来获得信息,这种方式的信息实时性差。服务器推送是一种服务器能主动向客户端推送实时信息的技术。本文提出了一种新的服务器推送技术,将其运用到交通监控平台上,实现交通信息实时发布的功能。

1 服务器推送技术

Web 系统的信息传递是通过 HTTP 协议进行的,HTTP 是一个无状态的单向协议,传统的 Web 通信很难做到实时性^[1]。服务器推送技术可以实现 Web 服务器与客户端之间的实时通信。服务器推送技术在现实应用中的解决方案大致可分为两类:一类是需要用户在浏览器安装插件,如 Flash, ActiveX, Applet, 插件技术要求用户安装相应的插件,这会影响到用户的体验效果。另一类无需安装插件^[2],无插件技术又可以分为客户端拉拽技术和基于 HTTP 长连接技术。客户端拉拽技术的主要实现模式为轮询模式;基于 HTTP 长连接技术又被称为 comet 技术,它的主要实现模式为长轮询模式和基于 iframe 流模式^[3]。

轮询模式使用 AJAX 技术周期性地向服务器发出调用

请求,服务器的数据就会不断被“请求”到客户端,从而实现服务器数据即时更新到客户端的目的^[4]。长轮询模式在目前 Comet 技术中应用广泛,浏览器首先发送一个 HTTP 请求,服务器端接收到该请求后,如果服务器端有数据则立即响应,如果服务器端没有数据产生则保持该 HTTP 请求处于等待状态,直到服务器产生数据并对该请求做出响应,也就是服务器端会保持长连接直到有数据传递或者超时才返回^[5]。浏览器接收到并处理完服务器端的响应后立即发送另一个新请求给服务器,这样一直循环下去直到双方通信结束。流模式能够提供较好的数据实时性,当浏览器向服务器发送一个请求时,服务器会一直保持该请求的连接,只要服务器端有新的数据产生就立即通过该连接发送给浏览器,实现数据的实时推送^[6]。流模式一般通过隐藏的 iframe 来请求数据,它将 iframe 的 SRC 属性设为对一个长连接的请求,之后服务器端就能够源源不断地往客户端发送实时数据,大大减少发送到服务器的请求次数,从而避免了与设置服务器连接相关的开销和延时^[7]。

2 现有推送方式的不足

3 种推送方法各有特点,有着不同的适用环境。轮询模式适合应用简单、请求事件的周期固定、事件频率可预知,同时对实时性要求不高的应用,如人人网;长轮询模式适合事件发生频率和实时性一般的应用,它能够接受事件发生频率较

到稿日期:2013-07-09 返修日期:2013-10-25 本文受上海市科技创新行动计划(10DJ1400300),国家国际合作专项(2012DFG11580),国家自然科学基金项目(61003221)资助。

蒋乾悦(1989—),男,硕士生,主要研究方向为 Web 信息服务,E-mail:jqy.de@qq.com;张亚英(1975—),女,副教授,硕士生导师,主要研究方向为协调技术、分布计算。

小的情况,同时能够保证较高的数据实时性,比如网站监控服务平台;流模式适合事件发生频率高并且对实时性要求较高的情况,但是由于流模式一直保持与服务器的连接,因此消耗的内存等系统资源相对比较多,比如黄金、股票报价网站。

对于事件发生频率波动较大、用户对推送要求的差异也较大的系统来说,以上3种单一固定的推送方案都不能很好地完成推送工作。比如一个实时交通监控平台,它会把用户关注路段的交通信息推送给当前用户。道路交通事故发生的频率是不可预知的,而且不同的用户对信息推送的实时性要求差异也很大,如交通管理部门需要在第一时间知道事故信息,而普通出行者对某一路段的事故信息不一定很敏感,因此现有的推送模型不能很好地满足交通监控平台的需求。为了更好地实现交通监控平台的信息实时发布功能,本文提出基于模糊决策的复合式轮询推送技术。

3 基于模糊综合决策模型的复合式轮询

当前一些对传统推送方法的改进研究是根据实际情况通过一些调度决策算法在两种常用推送技术(轮询,长轮询)中选择其中一个作为当前推送方案。由于这些改进的动态推送方法中可供选择的推送方法只有轮询和长轮询,并且轮询和长轮询的实时性和服务器资源利用率差别很大,因此上述改进算法较难在保证服务器性能充分发挥的同时根据实际用户情况提供合适的实时性。不同于其它对传统推送方法的改进,本文提出了一种新的推送方案——基于模糊综合评价模型的复合式轮询,此方案能根据不同实际条件动态调整推送方案的实时性,且实时性的高低能均匀分布在轮询和长轮询之间。

基于模糊综合评价模型的复合式轮询是一种能根据不同用户、不同登入时间、不同关注路段等情况进行动态调整的推送方法。它结合了轮询与长轮询的特点,并且运用模糊综合评价模型实现了推送方案的灵活性。

3.1 模糊综合评价模型

运用多个指标对多个参评单位进行评价的方法,称为多变量综合评价方法(又称综合评价方法)^[8]。其基本思想是将多个指标转化为一个能够反映综合情况的指标来进行评价。

模糊综合评价法是一种基于模糊数学的综合评标方法。该综合评价法根据模糊数学的隶属度理论把定性评价转化为定量评价,用模糊数学对受到多种因素制约的事物或对象做出一个总体的评价^[9]。在模糊综合评价模型中的元素有:因素集 E 、评语集 R 、因素权重分配向量 W 、评语权重分配向量 P 、综合评判变换矩阵 M 、综合决策 D ^[10]。

模糊综合评价模型的计算步骤如图1所示。

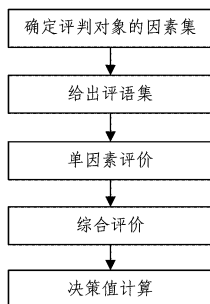


图1 模糊综合评价模型的计算步骤

(1)确定评判对象的因素集 $E = \{e_1, e_2, e_3, \dots, e_n\}$ 。如评判交通监控平台的信息推送,可记为 $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$,其中 $e_1, e_2, e_3, e_4, e_5, e_6$ 分别表示用户权限、用户登入时间、用户关注路口数量、路口、用户登入数量、CPU 占用率。

(2)给出评语集 $R = \{r_1, r_2, r_3, \dots, r_m\}$ 。如评判交通监控平台的信息推送,可记为 $R = \{r_1, r_2, r_3, r_4\}$,其中 r_1, r_2, r_3, r_4 分别表示非常重要、重要、一般、不重要。

(3)单因素评价。建立一个从 E 到 R 的模糊映射: $f: E \rightarrow F(R), x_i \mapsto \frac{a_{i1}}{v_1} + \frac{a_{i2}}{v_2} + \dots + \frac{a_{im}}{v_m}$ 。其中, $0 \leq a_{ij} \leq 1, i = 1, 2, \dots, n, j = 1, 2, \dots, m$ 。由 f 诱导出模糊关系 M_f ,用矩阵 M 表示,

$$M = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix}$$

称 M 为综合评判变换矩阵。其中,

(4)综合评价, $D = W \circ M$,其结果为单行 n 列的矩阵。其中, W 为各因素重要程度的集合,与因素集 E 相对应。例如在评判交通监控平台的信息推送模型中 $W = (0.35, 0.15, 0.15, 0.15, 0.05, 0.15)$ 。 \circ 为模糊算子 $M(+, \cdot)$,该模糊算子不仅考虑了所有因素的影响,而且保留了单因素的评价信息,全面地考虑了各个因素的影响和各个因素的评价结果。算子中 $+$ 表示普通加法, \cdot 表示普通乘法。即: $B_j = \sum_{i=1}^m a_i r_{ij} (j = 1, 2, \dots, n)$ 。

(5)决策值计算, $T = DP^T$,其结果为一个自然数。结果的大小可直接体现出被决策事物的重要程度。

3.2 复合式轮询

复合式轮询也是基于 HTTP 长连接的推送方式,它在长轮询的基础上进行了改进。长轮询是客户端与服务器的长连接一旦断开后又立即重新建立连接,而复合式轮询是断开长连接后客户端并不立即重新建立连接而是每次都会等待一段时间。图2为复合式轮询与长轮询的对比。等待时间 t 的取值是根据不同用户的实际情况通过模糊综合决策模型进行计算的。

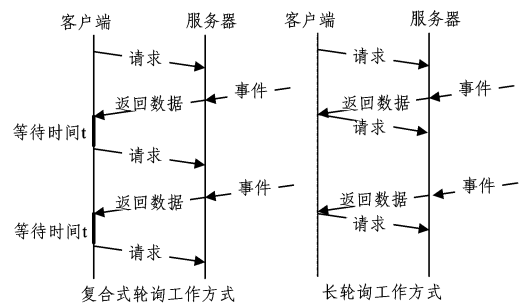


图2 复合式轮询与长轮询的对比

复合式轮询可以减少大量客户端连接对服务器造成的压力。由于在大量的客户端中每个用户对信息实时性的要求不同,比如交通管理部门用户在交通高峰时段登入系统时的实时性要求要远高于普通用户在交通低峰时段登入时的实时性要求。复合式轮询是一种能根据各种外界因素做出动态调整的推送方案。

4 交通监控平台信息发布

交通监控平台是一个向交通管理部门和普通用户提供道路交通信息的服务平台。它主要有交通路况查询、交通直播、

违章事故告警、事故查询等功能。本文将基于模糊综合评价模型的复合式轮询运用于违章事故告警部分,以实现事故实时告警。

在本服务平台中有两种不同权限的用户:政府用户和普通用户,政府用户对信息推送的实时性要求高于普通用户。每个用户登入时都会选择自己关注的路口,同时不同用户关注着不同数量的路口,同时每个路口的交通流量也不同。基于以上几点再结合服务器性能,得出影响复合式轮询等待时间 t 的因素有 6 个,分别为用户权限、用户登入时间、用户关注路口数量、路口流量、用户登入数量以及服务器 CPU 占用率。

等待时间 t 的计算步骤:

第 1 步 确定评语集和影响因素集的权重以及各因素在不同的条件下的评语模糊向量。在本系统中影响因素集合为(用户权限,用户登入时间,用户关注路口数量,路口,用户登入数量,CPU 占用率)。影响因素的权重是根据各影响因素的重要程度由专家组分别打分,最后统计专家分数并归一化而确定的,本系统中因素集的权重为(0.35,0.25,0.15,0.15,0.05,0.05)。权重越高,因素的变化对最后决策结果的影响越大。评语集为(非常重要,重要,一般,不重要),评语集的权重为(10,7,3,0)。综合决策值的取值范围在评语集权重的最大值和最小值之间,所以在平台中决策值的值域为[0,10]。

权重的取值在整个评价系统中很重要,它们是否符合实际需求情况直接影响到评价结果的好坏。权重的取值主要通过客观数据统计和专家组实践经验确定。例如通过统计上海市某一天的交通违章信息来确定用户登入时间和路口评语模糊向量。统计信息中包括车辆违章的时间的分布情况。表 1 所列全天车辆违章数在不同时间段的分布比例。

表 1 车辆违章数在不同时间段的分布比例

时间段	分布比例
7:01-9:30	12%
9:31-16:30	52%
16:31-19:30	8%
19:31-23:30	16%
23:31-7:00	12%

本文用 (a, d, c, b) 这 4 个变量来表示各时段的评语模糊向量,如表 2 所列。

表 2 各时段的评语模糊向量

用户登入时间	非常重要	重要	一般	不重要
7:01-9:30	a_1	b_1	c_1	d_1
9:31-16:30	a_2	b_2	c_2	d_2
16:31-19:30	a_3	b_3	c_3	d_3
19:31-23:30	a_4	b_4	c_4	d_4
23:31-7:00	a_5	b_5	c_5	d_5

由表 1 可得各时间段权重的比值关系。各时段权重的表

示方法为评语模糊向量乘以评语集: $(a, b, c, d) \begin{pmatrix} 10 \\ 7 \\ 3 \\ 0 \end{pmatrix} = 10a +$

$7b + 3c$ 。列出 5 组不等式:

$$\begin{cases} 10a_1 + 7b_1 + 3c_1 = 1.2e \\ a_1 + b_1 + c_1 + d_1 = 1 \\ a_1, b_1, c_1, d_1 > 0 \end{cases}$$

$$\begin{cases} 10a_2 + 7b_2 + 3c_2 = 5.2e \\ a_2 + b_2 + c_2 + d_2 = 1 \\ a_2, b_2, c_2, d_2 > 0 \end{cases}$$

$$\begin{cases} 10a_3 + 7b_3 + 3c_3 = 0.8e \\ a_3 + b_3 + c_3 + d_3 = 1 \\ a_3, b_3, c_3, d_3 > 0 \end{cases}$$

$$\begin{cases} 10a_4 + 7b_4 + 3c_4 = 1.6e \\ a_4 + b_4 + c_4 + d_4 = 1 \\ a_4, b_4, c_4, d_4 > 0 \end{cases}$$

$$\begin{cases} 10a_5 + 7b_5 + 3c_5 = 1.2e \\ a_5 + b_5 + c_5 + d_5 = 1 \\ a_5, b_5, c_5, d_5 > 0 \end{cases}$$

为了使最后的决策值能取到理论最高值“10”,令每个影响因素的最高权重为(1,0,0,0)。因为在“9:31-16:30”这个时间段交通违章数最高,所以它的评语模糊向量为(1,0,0,0)。将(1,0,0,0)代入第 2 个不等式组,可以得到 $a_2 = 1, b_2 = 0, c_2 = 0, d_2 = 0, e = 1.9$ 。把 e 的值代入其他不等式组可求得 $a_1 = 0.03, b_1 = 0.2, c_1 = 0.19$,通过这个方法可以得到所有的用户登入时间和路口评语模糊向量。而其他影响因素的评语模糊向量主要是根据实际需求情况和经验确定的。

第 2 步 给出每个影响因素的模糊映射

$$f: E \rightarrow \&(R) e_i \mapsto f(e_i) = (a_{i1}, a_{i2}, a_{i3}, \dots, a_{in})$$

其中, $f(e_i)$ 是有关因素 e_i 的评语模糊向量,是对 e_i 的一个评价, a_{ij} 是指 e_i 具有评语 r_i 的相关程度,也就是评语 r_i 在所有评语中所占的权重, $0 \leq a_{ij} \leq 1$ ($\sum_{j=1}^m a_{ij} = 1; i = 1, 2, 3, \dots, n; j = 1, 2, 3, \dots, m$)。比如当普通用户登入时用户这个因素的评语模糊向量为(0.02,0.1,0.8,0.08)。

第 3 步 求出综合评价变换矩阵。将上一步中得到的各因素的评语模糊权重向量写成矩阵关系,得到模糊关系矩阵 M ,称为综合评价变换矩阵, $M = (a_{ij})_{n \times m}$ 。它是某类事物的所有因素的各个评语相关性的综合表示。比如某一次的登入环境为:普通用户;8:00 登入;用户关注了 6 个路口;权重最高的路口在郊区;当前登入用户数为 16;当前服务器 CPU 占用率为 6%。查询表 1 就可以得到综合评价变换矩阵 M :

$$M = \begin{pmatrix} 0.02 & 0.5 & 0.4 & 0.08 \\ 0.05 & 0.2 & 0.7 & 0.05 \\ 0.02 & 0.3 & 0.6 & 0.08 \\ 0.3 & 0.6 & 0.08 & 0.02 \\ 1 & 0 & 0 & 0 \\ 0.6 & 0.3 & 0.05 & 0.05 \end{pmatrix}$$

第 4 步 综合评判。通过综合评判变换矩阵 M 将因素集 E 上的模糊集 $W = (w_1, w_2, w_3, \dots, w_n)$ 转换为评语集 R 上的模糊集,即综合决策 D 。

$$D = WM = (d_1, d_2, d_3, \dots, d_n)$$

其中, $d_i = \sum_{k=1}^m w_k \times m_{ki}, i = 1, 2, 3, \dots, m$ 。

因为在在本系统中 $W = (0.35 \ 0.25 \ 0.15 \ 0.15 \ 0.05 \ 0.05)$,所以根据上一步求出的 M 可以求出 D 。 $D = WM = (0.2025 \ 0.385 \ 0.3545 \ 0.058)$

第 5 步 求综合决策值。综合决策值即某类事物的最终优先等级,其中 $P = (P_1, P_2, P_3, \dots, P_m)$,代表评语集 R 的一个权重分配向量,其转置矩阵为 P^T ,将 D 与 P^T 按照普通的

矩阵乘法所得到的值为最终的综合决策值 T 。第 1 步确定了评语集及其权重, $P=(10\ 7\ 3\ 0)$ 。根据上步求出的 D 可以求出 T :

$$T = DP^T = (0.1425\ 0.445\ 0.3545\ 0.058) \begin{pmatrix} 10 \\ 7 \\ 3 \\ 0 \end{pmatrix} = 5.6$$

决策值 T 越大重要性越高, 等待时间 t 就越小。

第 6 步 求等待时间 t 。首先计算出在最坏情况下(各影响因素的评语为最差)的综合决策值 T_1 , 再计算出在最好情况下的综合决策值 T_2 。确定综合决策值的范围在 $[T_1, T_2]$ 之间。然后根据给定等待时间 t 的取值范围 $[t_1, t_2]$ 求出 t 。我们可以计算出 T 的取值范围为 $[10, 0]$, 但在实例中分别取各影响因素的最高权重和最低权重, 计算出 T 的实际取值范围为 $[2.9975, 10]$ 。在本平台中给定等待时间 t 的取值范围为 $[0, 60]$, 单位为 s。由于决策值 T 越大等待时间 t 就应该越小, 并且 t 要随着 T 均匀分布, 因此可以得到 $t(T)$ 函数: $t = 85.73 - 8.57T, T \in [2.9975, 10]$ 。根据上一步求出的 T 可以得到等待时间 t 的取值为 37.7s。

5 性能测试与分析

其他针对传统推送方法的改进研究是根据某种决策方法, 结合当前实际情况能够动态选择是用轮询方案推送还是长轮询方案推送。由于推送方法要么是轮询要么是长轮询, 而且轮询的实时性远不如长轮询的实时性, 因此其他针对传统推送的改进方案较难在保证服务器性能充分发挥的情况下根据实际条件提供合适的实时性。不同于其他针对传统推送方法的改进, 本文提出了一种新的推送方案, 此方案能根据不同实际条件动态调整推送的实时性, 且实时性的高低能均匀分布在轮询和长轮询之间。

本文从用户和服务器两个方面测试和分析基于模糊综合决策模型的复合式轮询的性能。

表 3 列举了在不同用户的登入情况下 t 的取值。

表 3 不同用户的登入情况下 t 的取值

用户 权限	登入 时间	用户关注 路口数量	权重最 高路口	当前用户 登入数量	服务器 CPU 占用率	t (s)
政府	10:00	13	中环内	15	10%	0
普通	10:00	13	中环内	15	10%	10.78
政府	8:00	13	中环内	15	10%	4.95
普通	22:00	4	外环外	9	6%	38.62
普通	14:00	10	内环内	50	28%	16.23

表 3 的第一行为各个因素取最高权重计算出来的 t 最小值。表中第二行除了用户级别与第一行不同, 其他条件都相同, 它计算出来的 t 值比第一行多 11s。因为政府用户对信息的实时性要求远高于普通用户对实时性的要求, 所以在相同条件下普通用户的轮询等待时间高于政府用户是合理的。表中第三行除了用户登入时间与第一行不同, 其他条件都相同, 它计算出来的 t 值为 5s。因为 8 点所在的时间段交通违章信息要少于 10 点所在的时间段, 所以第三行的取值也是合理的。通过表 3 的比较和分析表明, 模糊综合决策模型可以根据不同用户情况计算出能满足用户实时性要求并且合理的轮询等待时间 t 。

在服务器方面, 本文通过实验分析基于模糊综合模型的复合式轮询和现有的 3 种推送方式在 CPU 占用率、内存占用率实时性、数据丢失率、长连接时间占用率这 5 个性质上的区别。其中长连接时间占用率是指客户端与服务器保持长连接的时间占总时间的比值, 其值越低服务器资源利用效率就越高。

本文在交通监控平台中分别用长轮询、轮询、流模式和基于模糊综合模型的复合式轮询来实现实时事故信息发布。然后本文模拟 5 个政府用户和 5 个普通用户在事故多发时间段登入 4 个不同推送方案的平台, 并运用模拟违章事故信息发生程序随机产生 500 个违章信息。各推送方式的数据如表 4 所列。

表 4 各推送方式的测试数据

	复合式轮询	长轮询	轮询	流模式
平均 CPU 占用率	69%	85%	41%	89%
平均内存占用率	64%	70%	51%	77%
平均延迟时间	26s	14s	55s	1s
长连接时间占用率	75%	90%	1%	99%
数据丢失率	1%	1%	2%	0%

分析表 4 中的数据并结合各推送方式的自身特点, 总结各推送方式特性, 如表 5 所列。

表 5 各推送方式特性

	复合式轮询	长轮询	轮询	流模式
服务器资源占用率	一般	低	较高	高
服务器资源利用率	一般	低	高	低
实时性	一般	低	较高	高
数据丢失率	一般	一般	高	低
工作方式	异步	异步	同步	异步
灵活性	高	低	低	低

从表 5 可以看出, 复合式轮询与传统推送方式相比最大的优势是具有其他推送方式没有的灵活性, 并且其余各项指标均为中等值。复合式轮询的实时性高于轮询, 低于长轮询和流模式, 但复合式轮询的服务器资源利用率要高于长轮询和流模式。

结束语 本文提出的基于模糊综合决策模型的复合式轮询, 通过模糊理论将长轮询方案和轮询方案结合, 弥补了单一推送方案缺乏灵活性的不足。本文还将复合式轮询运用于交通监控平台中以实现信息的实时发布, 并进行实验、比较、分析, 结果表明本文推送方案既能满足不同用户的实时性需求又能节省服务器资源, 适合于用户数量大且对实时性需求变化大的平台。

参考文献

- [1] Gustitus C. The Push Is on; What Push Technology Means to The Special Librarian[J]. Information Outlook, 1998, 2(1): 21-24
- [2] 孙清国, 朱玮. Web 应用中的服务器推送技术研究综述[J]. 计算机系统应用, 2008(11): 116-120
- [3] Kendall J E, Kendall K E. Information Delivery Systems; An Exploration of Web Pull and Push Technology[J]. Communication of the AIS, 1999, 1(4): 16-18
- [4] 晏勇智. 基于 HTTP 长连接的服务器推送技术研究[D]. 长沙: 湖南大学, 2010: 10-11
- [5] Bozdog E, Mesbah A, Van Deursen A. A comparison of push and

pull techniques for ajax[C]//9th IEEE International Workshop on Web Site Evolution, 2007(WSE 2007). IEEE, 2007;15-22

- [6] 薛真真. 基于服务器推送和事件流处理技术的实时 Web 系统研究[D]. 杭州, 浙江大学, 2008;22-23
- [7] Zhou Quan, Bian Rui-xiang, Pan Yue-hun. Design of Electric Power Web System Based on Comet[C]//Second International Conference on Intelligent Computation Technology and Auto-

mation, 2009(ICICTA'09). IEEE, 2009, 3;42-45

- [8] 王海军, 张德礼. 多因素综合评价法划分征地区片方法研究[J]. 华中师范大学学报, 2006, 40(4);614-615
- [9] 李玲玲, 李凤强, 王成山. 基于模糊综合决策的配电网单相接地故障选线[J]. 电力系统及其自动化学报, 2011, 23(5);43-46
- [10] 段海彦. 基于动态调度机制的服务器推送技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2011;23-24

(上接第 58 页)

含上跳变和下跳变);第 4 列为用已知定时的测试产生方法得到的 TDF 覆盖率;第 5、6 列是在超速测试中采用单通路敏化标准并选用高测试质量优先的通路选择策略得到的 TDF 覆盖率以及 CPU 运行时间。最后两列是采用不同敏化方法得到的 TDF 覆盖率以及 CPU 运行时间,在通路选择上基本采取文献[11]的方法。在测试产生时间方面,已知定时的测试

产生方法将近传统测试产生时间的 20 倍,与后面两种方法不具有可比性。从实验数据可以看出,用不同敏化方法进行的测试虽然在测试时间方面相对文献[11]有所增加(主要是用非强健测试增加了测试产生时间),但在 TDF 平均覆盖率方面是 3 种方法中最高的,达到 86.0%,即采用不同敏化方法能以较小的时间代价提高小时延缺陷的 TDF 覆盖率。

表 3 实验数据对比

电路名称	PDF 总数	TDF 总数	文献[13]中,已知定时的测试产生的 TDF 覆盖率	文献[11]中采用单通路敏化的 TDF 覆盖率		用不同敏化方法的 TDF 覆盖率	
				TDF 覆盖率	CPU(s)	TDF 覆盖率	CPU(s)
S1423	89452	3004	85.82%	94.44%	4.00	95.02%	5.12
S5378	28092	11046	84.84%	93.13	5.28	94.83%	5.88
S9234	489708	18968	81.30%	68.41%	20.4	75.60%	26.35
S13207	2690738	27938	79.62%	83.89%	38.59	85.84%	52.20
S15850	329467092	33062	70.26%	79.21%	212	79.08%	270
S35932	394282	75320	82.10%	61.22%	576	74.32	615
S38417	2783158	80162	98.04%	89.70%	371.6	93.96%	442
S38584	2161446	80324	83.86%	88.04%	352	89.35%	427
平均			83.23%	82.27%		86.0%	

结束语 采用不同敏化标准进行小时延检测有以下优点:(1)采用高效的可达通路选择方法以及 fan 算法里面的相关技术,使得测试产生效率高。(2)采用超速测试,并对路径进行分组,选取通路时延接近于给定测试时钟周期的通路组进行测试,能保证测试质量。(3)根据路径长短以及被敏化的可能性高低来选取合适的敏化标准,能提高小时延的 TDF 覆盖率。

在本文工作基础上,可以朝下面几个方向继续深化:(1)研究更接近电路真实时延的统计定时分析方法;(2)在统计定时分析方法基础上,研究针对特定时延影响因素的故障模型和测试产生方法,以提高测试质量;(3)研究高效率测试产生算法。

参 考 文 献

- [1] Breuer M A. The effect of races, delays and delay faults on test generation[J]. IEEE Transactions on Computers, 1974, 22(10); 1078-1092
- [2] <http://www.itrs.net/Links/2008ITRS/Home2008.htm>
- [3] Qiu W, Wang J, Walker D M H, et al. K longest paths per gate (KLPG) test generation for scan-based sequential circuits[C]// Proc. Int. Test Conf. (ITC). 2004;223-231
- [4] Sato Y, Hamada S, Maeda T, et al. Invisible delay quality-SDQL model lights up what could not be seen[C]// Proc. Int. Test Conf. (ITC). 2005;47
- [5] Kajihara S, Morishima S, Takuma A, et al. A framework of high-

- quality transition fault ATPG for scan circuits[C]//Proc. Int. Test Conf. (ITC). 2006
- [6] Shao Y, Pomeranz I, Reddy S M. On generating high quality tests for transition faults[C]//Proc. Asian Test Symp(ATS). 2002;1-8
- [7] Lin X, Tsai K, Wang C, et al. Timing-aware ATPG for high quality at-speed testing of small delay defects[C]//Proc. Asian Test Symp. (ATS). 2006;139-146
- [8] 王杰,李华伟,梁华国. 针对小时延缺陷的时延测试方法综述[C]//第十三届全国容错计算学术会议. 2009;383-390
- [9] Lee B N, Wang L C, Abadir M S. Reducing pattern delay variations for screening frequency dependent defects[C]//Proc. VLSI Test Symp. (VTS). 2005;153-160
- [10] Ahmed N, Tehranipoor M. A novel faster-than-at-speed transition-delay test method considering IR-drop effects[J]. IEEE Trans. Comput.-Aided Design (CAD) Integr. Circuits Syst., 2009, 28(10);1573-1582
- [11] Fu Xiang, Li Hua-wei, Li Xiao-wei. Testable Path Selection and Grouping for Faster Than At-Speed Testing[J]. IEEE Trans. on VLSI Systems, 2012, 20(2);236-247
- [12] Fujiwara H, Shimono T. On the Acceleration of Test Pattern Algorithms[J]. IEEE Transactions on Computers, 1983, 32(12);1137-1144
- [13] Kajihara S, Morishima S, Takuma A, et al. A framework of high-quality transition fault ATPG for scan circuits, [C]// Proceedings IEEE of International Test Conference(ITC). 2006;2