

# 一种新型的 Web 应用 可靠性评估工作量计算方法

崔霞<sup>1</sup>, 袁佳琳<sup>2</sup>, 高建华<sup>2</sup>

(1. 华东师范大学软件学院, 上海 200062; 2. 上海师范大学信息与机电工程学院, 上海 200234)

**摘要:** Web 应用系统已经应用到各个领域, 研究 Web 应用的可靠性日益重要. 首先定义 Web 应用失效, 引出了 Web 应用可靠性的概念; 通过分析 IIS 日志数据, 收集相应的用户操作和信息传送失效数据, 从 Web 应用软件系统的角度探讨了基于日志的 Web 应用可靠性评估的可行性. 基于网站用户群使用特征的规律性, 提出了一种新型工作量计算方法, 该方法首先运用平均加权技术进行无单位化处理, 再设置目标函数经优化得到计算新型工作量的权量. 最后通过实验验证了新型工作量的有效性. 实验结果表明, 采用新型工作量计算 Web 应用的可靠度稳定性较好, 可信性较高.

**关键词:** Web 应用; 工作量; 可靠性; 优化

**中图分类号:** TP 39   **文献标志码:** A   **文章编号:** 1000-5137(2015)01-0105-06

## 0 引言

计算机技术已全面渗透到社会活动的各个领域, 不同领域相互之间的交叉, 衍生出了 Web 应用, 它为大规模的用户群提供了跨平台访问互联网资源的各种途径. Web 应用系统已经完成由“胖客户端”模式向“瘦客户端”模式转变, 目前更多的信息系统使用 WWW 承载 B/S 结构<sup>[1]</sup>. 人们对 Web 应用的依赖度日益增强, 确保其可靠性变得愈加重要.

传统软件的可靠性度量方法都是基于软件内部故障. 而 Web 应用是由一系列 Web 页面及部件构成的复杂的、多层次的、异构的系统. 它的执行需要浏览器、网络、Web 服务器、应用服务器、数据库服务器等共同协调<sup>[2]</sup>. 这些特点导致 Web 失效因素的多样性, 不仅包含软件内部故障, 更包含用户使用、网络环境等诸多因素; 而且一种失效现象可由多种可能原因导致.

文献[3]针对 Web 应用可靠性评估, 首次提出基于 Apache 日志的工作量提取方法, 着重探讨了分别使用点击数、下载字节数、用户数以及用户会话数作为工作量指标, 度量 Web 软件可靠性的可行性. 上述方法虽然避免了以时间作为可靠性模型输入所带来的巨大误差, 但是其强调的 Web 应用失效重点仅在于源内容的失效, 而忽略了由于网络环境故障等因素所产生的影响.

本文作者从软件系统角度探讨基于日志的 Web 应用可靠性评估的可行性, 包括失效类型及原因分析、基于 IIS 日志的数据收集, 以及新型的工作量计算方法. 将新的工作量指标代入 Nelson 模型, 评估 Web 应用的可靠性, 并衡量可靠性度量的精度.

章节安排如下: 第 1 节是 Web 应用的可靠性, 主要描述 Web 应用失效定义以及失效形式. 第 2 节是基于 IIS 日志的数据收集, 描述 IIS 访问日志和错误日志的采集过程和与工作量相关数据处理过程. 第 3

收稿日期: 2014-12-20

基金项目: 国家自然科学基金资助项目(61073163); 上海市企业自主创新专项资金项目(沪 CXY-2013-88)

通信作者: 崔霞, 中国上海市中山北路 3663 号, 华东师范大学软件学院, 邮编: 200062, E-mail: cuixia\_O@126.com

节是核心部分,新型工作量计算方法.描述所提出的新型工作量的计算方法.第4节是实验分析,根据所采集的日志数据,对新型工作量的准确性根据 RSE 值进行评价.第5节是结语.

## 1 Web 应用的可靠性

将 Web 应用失效定义为在用户正确操作下,无法完成文档、多媒体以及应用计算结果等信息传送的事件.导致 Web 信息传送失败的原因主要有以下 3 个方面:1) 用户操作错误;2) 网络环境故障;3) 源内容失效及网页错误.其中,源内容失效通常指服务器端信息的丢失、传输阻碍及程序运行失败;网络环境故障通常指由于客户端浏览器的专有特性,网络设备的运行局限等各种因素所导致的不稳定、网路阻碍等现象.显而易见,能导致所定义的 Web 应用失效的原因就是上面所提及的 2) 和 3).而 1) 用户操作错误与研究内容无关.

Web 应用更新频率较高,访问量波动较大等特点促使用户更关心信息的正确获取.软件可靠性模型是软件可靠性定量分析的技术基础.将 Web 应用可靠性定义为避免 Web 应用失效,无故障完成信息传送的概率.由于 Web 应用中确切的时间是很难获得的,工作量概念的提出避免了可靠性模型的输入基于 CPU 运行时间,事务处理时间等指标的尴尬<sup>[3]</sup>.

Web 日志文件详细记录了客户端和服务器的交互行为和错误类别.举例来说,如下错误日志内容:2010-12-25 02:22:49 111.186.89.52 1226 202.121.48.103 80 - - - - - Timer\_ConnectionIdle -,说明错误是于 2012 年 12 月 25 日 02:22:49 发生,错误类型为连接超时.由于 Web 应用失效因素错综复杂,往往很难区分失效的具体原因和出错对象,如何从 Web 日志中有效提取工作量数据和失效数据决定着可靠性评估的成败.

## 2 基于 IIS 日志的数据收集

IIS 是目前常用的服务器之一,IIS 访问日志和错误日志是作者的研究对象.IIS 访问日志主要用于记录用户包括搜索引擎蜘蛛对网站的访问行为,包括客户端访问时刻、来源 IP、客户端请求方式、访问路径及参数、协议状态代码、请求及返回字节大小等信息.IIS 错误日志主要用于记录 HTTP 应用程序接口错误,包括客户端访问时刻、来源 IP、谓动词、访问资源、协议状态代码、错误原因等信息.

需要注意的是,IIS 访问日志与错误日志是相互独立的,在访问日志中出现的错误信息,错误日志中并无记录与之对应.IIS 访问日志是服务器已经响应客户端的请求,并已经发生请求和返回信息传送的记录;而 IIS 错误日志则是由于网络环境故障、服务器和客户端错误等原因导致的信息传送中止的记录.

虽然很难从 IIS 访问和错误日志中严格区分信息传送失败的具体原因和出错对象,但通过恰当的处理可以较为准确地划分出所定义的 Web 应用失效次数和用户访问的信息.IIS 访问日志和错误日志记录了 Web 应用系统各个环节软件的响应,而与硬件设备的损坏无关,这正是从 Web 应用软件系统角度探讨可靠性评估的着眼点.作者通过 C# 语言编程实现了对 IIS 日志的数据收集.

### 2.1 工作量数据收集

参照文献[3]提出的 4 个工作量的候选参数,即,点击数、传输字节数、用户数、用户会话数,在第 3 节中对上述 4 个参数进行组合,产生一种新型的工作量的计算方法.这 4 个候选参数选取如下:

1) 点击数.每次点击代表了一次客户端与服务器的交互.由于 IIS 记录访问日志时以文件引用为一次记录,使得一次用户的访问请求记录了如 CSS 样式文件、图片文件和 JS 脚本文件的引用等多条记录.计算点击数时,是指去除了访问请求所附带的隐性资源、搜索引擎蜘蛛的访问,以及协议状态为 4xx 和 5xx 的错误访问记录后,剩余的访问日志条目总数即为点击数,单位“次”.

2) 传输字节数.在访问日志中可以获取一次点击的字节传输量,包括同一点击下多次文件引用所产生的客户端发送请求和服务器应答的总字节数,单位“字节”.文中排除了搜索引擎蜘蛛和错误访问记录中的传输字节数据.

3) 用户数. 一个唯一的 IP 对应一个用户. 用户数单位是“个”. 文中排除了搜索引擎蜘蛛和错误访问记录中的 IP 数据.

4) 用户会话数. 在一个明确的时间间隔内, 同一个 IP 地址上的所有操作被认为是一个会话, 超过这个时间间隔则产生一个新的会话, 用户会话数的单位是“次”. 文中排除了搜索引擎蜘蛛的 IP 数据, 并忽略了协议状态为 4xx 和 5xx 的点击记录.

### 2.2 失效数据收集

分别从 IIS 访问日志和错误日志中提取 Web 失效数据. 它反映了 Web 应用失效的来源, 体现了 Web 应用系统各个软件环节信息传送的基本运行状况.

每一次信息传送失败记为一次失效. IIS 访问日志中, 对应协议状态为 4xx 或者 5xx 的每一次点击记为一次失效; 大多数情况下, IIS 错误日志中, 错误原因“URL”、“Timer\_MinBytesPerSecond”、“Timer\_ConnectionIdle”、“Client\_Reset”和“Connection\_Dropped”是由用户错误操作或者硬件故障产生. 所以 IIS 错误日志中排除以上 5 种错误原因的每一次点击记为一次失效. Web 应用失效数的单位是“次”.

## 3 新型工作量计算方法

一般而言, 某网站的客户群体特点大致相同; 其客户的使用特征虽然在不断地波动, 但也具有一定的规律性. Web 日志中点击数、传输字节数、用户数以及用户会话数都存在信息不完整的现象. 基于以上分析, 采取了以下工作量的计算方法.

设某  $i$  天的点击数、传输字节数、用户数、用户会话数分别为  $H_i, B_i, U_i, S_i$ . 统计  $n_i$  天数的数据, 分别取平均, 得:

$$\bar{H} = \frac{\sum_1^{n_i} H_i}{n_i}; \bar{B} = \frac{\sum_1^{n_i} B_i}{n_i}; \bar{U} = \frac{\sum_1^{n_i} U_i}{n_i}; \bar{S} = \frac{\sum_1^{n_i} S_i}{n_i}. \quad (1)$$

再进行去单位化处理, 得:

$$h_i = \frac{H_i}{\bar{H}}; b_i = \frac{B_i}{\bar{B}}; u_i = \frac{U_i}{\bar{U}}; s_i = \frac{S_i}{\bar{S}}. \quad (2)$$

其中,  $h_i, b_i, u_i, s_i$  分别称为某  $i$  天的点击权重、传输字节权重、用户权重、用户会话权重.

定义某  $i$  天的工作量权重  $w_i$  为:

$$w_i = k_1 h_i + k_2 b_i + k_3 u_i + k_4 s_i \quad (k_1 + k_2 + k_3 + k_4 = 1, -1 \leq k_1, k_2, k_3, k_4 \leq 1). \quad (3)$$

其中,  $k_1, k_2, k_3, k_4$  的选取采用以下优化方法:

由 Nelson 模型, 某  $i$  天可靠度的估计值为:

$$r_i = 1 - \frac{n_i}{w_i \cdot \bar{W}}. \quad (4)$$

其中,  $w_i \cdot \bar{W}$  称为新型工作量, 若  $\bar{W}$  取  $\bar{H}$ , 则工作量表示为点击数; 同理, 若  $\bar{W}$  取  $\bar{B}$ , 则工作量表示为传输字节数.  $n_i$  是 Web 应用失效数. 由于文中 Web 应用失效的单位是“次”, 为保证可靠度计算的单位统一,  $\bar{W}$  取  $\bar{H}$ .

进一步计算  $n_i$  天可靠度的标准差, 定义为:

$$\chi_{n_i} = \sqrt{\frac{\sum_1^{n_i} (r_i - \bar{r})^2}{n_i - 1}}, \quad \left( \bar{r} = \frac{\sum_1^{n_i} r_i}{n_i} \right). \quad (5)$$

则,  $k_1, k_2, k_3, k_4$  的选取要求达到  $\chi_{n_i}$  最小. 文中运用 Matlab 软件的 fmincon 函数实现了  $k_1, k_2, k_3, k_4$  的优化选取.

为了说明可靠度评估的稳定性, 引入了 RSE 参数, 定义如下:

$$RSE = \frac{\chi_{n_i}}{\bar{r}}. \quad (6)$$

显然,  $RSE$  值越小, 可靠性评估越稳定, 可信性越高.

## 4 实验分析

本章将通过分析前 10d(2011.04.01 ~ 2011.04.10) 的 IIS 访问日志和错误日志数据, 获取点击数、传输字节数、用户数、用户会话数(时间间隔 20 min) 以及失效数. 利用第 3 节提出的优化方法, 选取  $k_1$ 、 $k_2$ 、 $k_3$ 、 $k_4$ . 再分析后 10d(2011.04.11 ~ 2011.04.20) 以及再后 10d(2011.04.21 ~ 2011.04.30) 的 IIS 访问日志和错误日志数据, 将前面优化得到的  $k_1$ 、 $k_2$ 、 $k_3$ 、 $k_4$  代入式(3), 计算新提出的工作量. 分别使用点击数和新型工作量, 代入 Nelson 模型(式(4)) 和  $RSE$  参数(式(6)), 计算可靠度和  $RSE$  值. 最后, 比较两种工作量计算可靠度的稳定性.

### 4.1 实验环境

本次实验使用一个社区网络运行商提供的网络业务办理平台. 社区网络用户可以在该平台上申请、办理、查询、咨询、报修各类网络业务. 本平台的 WEB 应用承载在 IIS6.0 上, 故本次实验使用 IIS6.0 自动记录的访问日志和错误日志. 要注意的是, IIS6.0 默认不记录每次用户请求时服务器接收和应答的字节数, 需要手动勾选 sc-bytes 和 cs-bytes 两个选项.

### 4.2 数据收集结果

分别选取了 2011 年 4 月 1 日至 10 日、11 日至 20 日、21 日至 30 日的 IIS 访问日志和错误日志作为实验数据, 从中获取点击数、传输字节数、用户数、用户会话数(时间间隔 20 min) 和失效数进行试验. 前 10d 的数据如表 1 所示; 后 10d 的数据如表 2 所示; 再后 10d 的数据如表 3 所示.

表 1 前 10d 的收集数据

天数	点击数	传输字节数	用户数	用户会话数	失效数
1	2406	65873163	120	161	302
2	1948	53129298	68	99	188
3	461	4080493	45	60	39
4	453	4253229	85	103	74
5	2154	59977444	131	176	234
6	2142	61641247	53	90	227
7	1942	56457453	77	121	260
8	1897	59384746	75	118	206
9	1982	59499650	94	165	228
10	313	124963	22	44	47

表 2 后 10d 的收集数据

天数	点击数	传输字节数	用户数	用户会话数	失效数
1	2181	71720430	92	165	281
2	1490	51366212	65	98	154
3	1656	51769976	67	100	207
4	1621	52566746	71	124	277
5	1616	56346749	91	144	326
6	181	646001	27	45	23
7	200	734394	48	73	52
8	2816	83028669	133	196	260
9	2320	76528687	139	173	379
10	2417	78105580	183	229	588

表 3 再后 10d 的收集数据

天数	点击数	传输字节数	用户数	用户会话数	失效数
1	1960	58238205	84	139	154
2	1539	42206898	49	77	98
3	495	1811612	21	39	11
4	468	618789	26	45	16
5	483	1338347	35	53	21
6	2196	71682668	106	163	170
7	2305	71140630	97	161	299
8	1898	59784539	72	108	167
9	461	170492	18	32	25
10	521	2377386	42	57	45

### 4.3 实验结果比较

通过前 10d 数据的优化计算,得到了系数  $k_1$ 、 $k_2$ 、 $k_3$ 、 $k_4$  值分别为 0.1220、0.4430、-0.4744、0.9094. 采用新型工作量前 10d 优化所得可靠度的  $RSE$  值为 0.0177;以点击数为工作量所得可靠度的  $RSE$  值为 0.0275. 代入上述系数  $k_1 \sim k_4$ ,采用新型工作量计算后 10d 可靠度的  $RSE$  值为 0.0586;若采用点击数为工作量,可靠度的  $RSE$  值为 0.0687. 采用新型工作量计算再后 10d 可靠度的  $RSE$  值为 0.0308;若采用点击数为工作量,可靠度的  $RSE$  值为 0.0335. 其柱状图如图 1 所示.

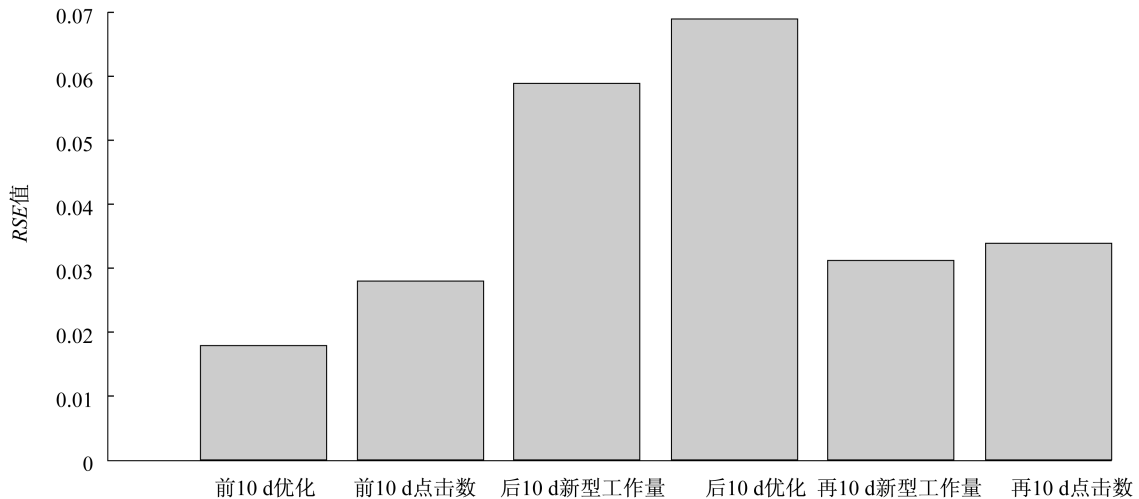


图 1 RSE 值比较

由实验结果数据可见,采用点击数作为工作量,3 组 10d 的可靠度  $RSE$  值分别为 0.0275、0.0687 和 0.0335,都是较差的. 而采用新型工作量经优化后的可靠度  $RSE$  值达到最小,即 0.0177;将优化后的系数  $k_1 \sim k_4$  代入后二组 10d 的数据,采用新型工作量计算可靠度的  $RSE$  值也分别达到 0.0586 和 0.0308,都比采用点击数为工作量所计算的  $RSE$  值低约 10%. 采用新型工作量所计算可靠度  $RSE$  值都相对较小,表明采用新型工作量来计算 Web 应用的可靠度稳定性较好,可信性较高.

## 5 结 语

本文作者通过定义 Web 应用失效为在用户正确操作下,无法完成文档、多媒体以及应用计算结果等信息传送的事件,来研究 Web 应用的可靠性. 通过分析 IIS 日志数据,收集相应的用户操作和信息传送失效数据. 从 Web 应用软件系统的角度探讨了基于日志的 Web 应用可靠性评估的可行性.

本文作者提出了一种新型工作量计算方法. 该方法首先从以前数日的日志数据中优化得到系数  $k_1 \sim k_4$ ,然后代入以后的日志数据,得到工作量权量  $w_i$  和新型工作量. 最后通过实验验证了新型工作量的有效性. 实验结果表明,采用新型工作量计算 Web 应用的可靠性较稳定,可信性较高.

## 参考文献:

- [1] LYU M R. Handbook of software reliability engineering[M]. New Yourk: McGraw-Hill, 1996.
- [2] ELBAUM S, ROTHERMEL G, KARRE S, et al. Leveraging user-session data to support web application testing[J]. IEEE Transactions on Software Engineering, 2005, 31(3): 187 – 201.
- [3] TIAN J, RUDRARAJU S, LI Z. Evaluating web software reliability based on workload and failure data extracted from server logs[J]. IEEE Transactions on Software Engineering, 2004, 30(11): 754 – 769.
- [4] YUAN J L, GAO J H. Measurement method of Web applicaiton reliability based on workload and error log[J]. Computer applications and software, 2012, 29(10): 44 – 47, 179.
- [5] NELSON E. Estimating software reliability from test data[J]. Microelectronics and Reliability, 1978, 17(1): 67 – 73.
- [6] W3C. Extended log file format[EB/OL]. (1996-03-23). [2014-11-20]. <http://www.w3.org/TR/WD-logfile.html>.
- [7] HSU C J, HUANG C Y. Optimal weighted combinational models for software reliability estimation and analysis[J]. IEEE Transactions on Reliability, 2014, 99: 1 – 19.

## A new measurement of workload in Web application reliability assessment

CUI Xia<sup>1</sup>, YUAN Jialin<sup>2</sup>, GAO Jianhua<sup>2</sup>

(1. School of Software, East China Normal University, Shanghai 200062, China;

2. Colloge of Information and Mechatronics, Shanghai Normal University, Shanghai 200234, China)

**Abstract:** Web application has been popular in various fields of social life. It becomes more and more important to study the reliability of Web application. In this paper the definition of Web application failure is firstly brought out, and then the definition of Web application reliability. By analyzing data in the IIS server logs and selecting corresponding usage and information delivery failure data, the paper study the feasibility of Web application reliability assessment from the perspective of Web software system based on IIS server logs. Because the usage for a Web site often has certain regularity, a new measurement of workload in Web application reliability assessment is raised. In this method, the unit is removed by weighted average technique; and the weights are assessed by setting objective function and optimization. Finally an experiment was raised for validation. The experiment result shows the assessment of Web application reliability base on the new workload is better.

**Key words:** Web application; workload; reliability; optimization

(责任编辑:包震宇)