

单点双信道网络测量系统的时钟误差概率模型

龚 俭^{1,2}, 徐加羚³

- (1. 江苏省计算机网络技术重点实验室, 南京 江苏 210096;
2. 中国教育和科研计算机网华东地区网络中心, 南京 江苏 210096;
3. 东南大学 计算机科学与工程系, 南京 江苏 210096)

摘要: 网络测量技术已成为当前互联网技术研究的热点之一. 时钟误差是网络测量中最主要的直接误差来源. 传统网络测量技术重视修正和提高时钟精度, 但忽视了测量中客观误差对结果的影响, 这种做法不适用于对高速网络的分布式协同测量中的误差进行定量描述. 基于此, 结合误差分析理论和网络被动测量的具体特点, 提出了针对单点双信道被动测量系统的一般性时钟误差概率模型, 引入误差不确定度来定量描述协调测量结果的精确度和合理性.

关键词: 计算机网络; 网络测量; 时钟误差分析; 概率模型

中图分类号: TP393.06 **文献标识码:** A

0 引言

随着网络技术的不断发展, 网络测量技术已成为互联网技术研究的热点之一. 其中, 基于被动抽样方式的测量技术和分布式协同测量体系结构的测量方式成为高速大规模 IP 网络测量技术方向的主要研究焦点. IETF 在 1998 年之前就成立了 IPPM、PSAMP、和 IPFIX^[1-3] 等工作组, 对相关的主题进行研究. 1998 年, 美国圣地亚哥大学也成立了 CAIDA^[4] (the Cooperative Association for Internet Data Analysis) 组织发起了全世界范围的互联网协同测量活动以及相关的学术讨论.

在已往的工作中, 网络测量试验的设计、试验数据的分析和相关探讨往往忽略测量过程中引入的误差. 与此同时, 研究用 NTP^[5] 方式进行网络钟同步研究作为一个独立的研究方向开始发展. 而对于时钟误差的研究主要侧重于如何提高时钟精度以及相关的时钟修正算法的研究和探讨. 因此时钟误差定量分析仍未被结合到测量研究之中.

随着网络技术的不断成熟, 网络发展呈现出高速和大规模的趋势. 当这一发展趋势使得网络测量技术的研究重点转移到对被动抽样测量和分布式协同测量的体系结构的研究, 测量中的误差对测量结

果的影响则越来越大. 首先, 网络传输速度的提高要求被动测量中涉及到的网络报文到达时间的记录精度不断提高数量级, 即测量中的相对误差对报文到达时间的影响作用也不断增加; 其次, 网络规模的无限延展要求网络协同测量方式, 而协同测量结果的可信度将受参与协同双方系统的精度以及测量任务的不同所引入的误差的极大影响. 因此, 如何对网络测量中的误差进行定量分析成为现代网络测量技术中不可避免的研究问题之一.

本文从实际测量系统的时钟偏差的产生现象入手, 分析了被动单点双信道系统中时钟偏差的概率分布特征, 进而提出了描述被动单点双信道测量系统的时钟误差概率模型, 以及如何用误差不确定度来描述测量结果误差和在协同测量中如何用不确定度来判定协同测量的可行性及结果的可信度.

1 测量器时钟误差的来源

目前基于通用硬件构架实现的测量器, 其时钟误差主要源于硬件时钟漂移和软件操作系统的非实时中断响应和任务调度切换^[6]:

$$T_m(t_0) = T'_h(t_0) + \Delta t_{OSIS} \quad (1)$$

其中: 下标 m, 表示测量; 下标 h, 表示硬件; 下标

收稿日期: 2005-08-09.

基金项目: 国家 973 计划课题(2003CB314803); 江苏省网络与信息安全重点实验室资助项目(BM2003201).

作者简介: 龚 俭(1957), 男, 博士, 教授, 博士生导师, E-mail: jgong@njnet.edu.cn; 徐加羚(1979), 男, 硕士生, E-mail: jlxxu@njnet.edu.cn.

ts, 表示切换.

由 $T'_h(t_0)$ 引起的时钟误差是目前对如何提高时钟精度研究的主要研究对象, 其分布特征和修正方法相对比较成熟. 该方向目前主要的研究成果包括以 NTP 采用的时钟修正算法为代表的软件修正方式和 Intel 提出的高精度时钟计数器 HTC 的硬件精度提高方式. 而网络测量系统 $\Delta t_{OS ts}$ 为本文的主要研究内容.

1.1 $\Delta t_{OS ts}$ 的试验概率分布现象

由于 $T'_h(t_0)$ 的存在, 无法直接对 $\Delta t_{OS ts}$ 进行测试. 必须首先消除 $T'_h(t_0)$ 对时钟偏差的影响以简化问题. 由于硬件时钟振荡误差相对于软件操作系统产生的切换延时非常短, 在短时间间隔 Δt 内, 相对于 $\Delta t_{OS ts}$ 而言,

$$T'_h(t_0) \approx T'_h(t_0 + \Delta t) \quad (2)$$

可以采用双信道同源报文采集方式来消除 $T'_h(t_0)$ 的影响. 试验中, 在单点测量系统上, 通过在两个网卡上接受来自同一数据源报文流来抵消 $T'_h(t_0)$ 的作用. 由式(1)可知, 在 Δt 时间内对同一报文在两块网卡上的时间戳分别为

$$\begin{aligned} T_{m1}(t_0) &= T'_h(t_0) + \Delta t_{1OS ts} \\ T_{m2}(t_0 + \Delta t) &= T'_h(t_0 + \Delta t) + \Delta t_{2OS ts} \end{aligned} \quad (3)$$

由(3)中两式相减, 消去 $T'_h(t_0)$ 影响后的表达式为 (即随机误差 $\Delta t_{OS ts}$ 的分布情况)

$$T_{m2}(t_0 + \Delta t) - T_{m1}(t_0) = \Delta t_{2OS ts} \text{ 或 } -\Delta t_{1OS ts} \quad (4)$$

1.2 测量器时钟误差分析试验

基于上述思路, 采用如图 1 所示方式连接. 将光纤 GE 以太网单向流量通过分光器产生相同的镜像流. 将镜像流分别接入测量器的两块采集千兆网卡, 在测量器上计算同一报文在两块网卡上的到达时间差.

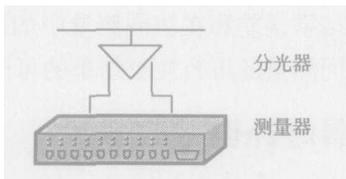


图 1 同源分光方式

Fig 1 Homologous optical splitting

被测系统采用基于 Ring-Buffer^[7] 零拷贝技术, 能支持 540 kbps 的高速流量测量, 是目前较为成熟和广泛采用的软件测量技术. 测量器采用 Intel 至强 2.4GB×2 的处理器和 2GB 物理内存的硬件配置; 操作系统平台采用 Linux 2.4.26 内核. 时间差的计算采用离线方式处理, 避免引入额外影响.

在 600 MB/s 实际流量下, 所得同报文时间差分布情况如图 2 所示. 图中延迟值基本集中在零点附近正 100 ns 的范围之内对称分布. 这与 Linux 内核的准实时调度方式相关.

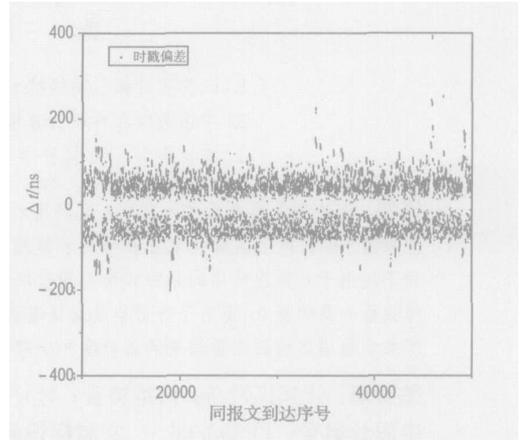


图 2 单点双信道同报文接收时间差

Fig 2 Homologous packet's absolute delay on single probe node's two channels

在多次试验中, 在系统轻负载情况下, 单点系统双信道同报文测试的时间差分布具有稳定性. 根据试验结果绘制时间差分布密度曲线如图 3 所示.

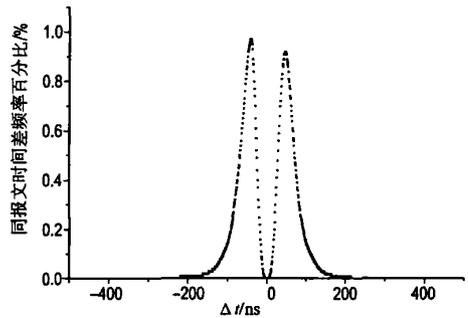


图 3 单点双信道同报文接收时间分布

Fig 3 Distribution of homologous packet's delay on single probe node's two channels

图 3 中, 在 0 点的两侧分布出现图形相近的双峰分布. 通过统计计算, 双峰的分布一致, 但是正负双峰有各自的均值和标准方差. 如表 1 所示.

表 1 双信道统计特征

Tab 1 Statistical features on twol

统计特征	信道 1	信道 2
均值	60.51	59.2
标准方差	35.81	35.08

两信道的时间差所反应出来的测量器时间误差分布具有相同或近似的概率分布. 其数字特征非常

接近, 但是通过 Wilcoxon 秩和检验证明两信道的分布为独立分布, 相关分布的概率 < 0.000 1. 这反映了在两个信道数据完全一样的情况下, 即流量强度一致, 双信道采集的时钟误差具有分布的相似性、统计特征的独立性和稳定性. 虽然两个信道在同一物理单点系统中, 但是其统计的特征相互独立. 而分布的相识性源于同一物理单点系统的双信道采集所采用同一的中断处理进程或线程调度机制. 而稳定性既是估算测量器时钟误差的前提要求, 也是系统在一定负载范围内的实际测试结果. 其原因主要在于: 一方面测量器不同于其他通用服务器, 测量任务进程受其他任务负面影响很小; 另一方面在硬件和操作系统允许工作压力范围内, 它们的设计要求系统运行具有稳定的概率特征, 包括中断切换时间、任务切换时间等.

2 单点双信道时间误差分布模型

2.1 误差分布模型的数学表达

在上述单点双信道同报文测试试验并不等同于测量器在实际运行中的环境, 在获得 Δt_{otis} 的实际分布情况的同时丢弃了几个实际测量中设计的关键因素对时间误差分布的影响. 首先, 试验中采用同报文方式; 同样的报文在两条信道上严格等比例分配, 这不同于实际情况中被测报文在信道上的分布概率与信道设置与测量任务需求有关. 其次, 试验中双信道上流量压力相同, 不同于实际中不同信道上的背景流量分布不均匀的情况. 因此不能将同报文延迟实验的结果直接应用在对实际运行中的测量器上.

实际测量环境中, 报文在信道上的分布, 信道的总体背景流量的比例关系以及测量任务的要求是除了测量器本身因素外的 3 个影响测量误差的主要因素.

考虑上述因素对时钟误差分布的影响, 建立如下的双信道采集的时钟误差模型公式:

$$\begin{aligned} P(x) &= P(A) \times P(x|A) + P(B) \times P(x|B) \\ P(x|A) &= P(x|A \& A_a) \times P(A_a) + \\ &\quad P(x|A \& A_d) \times P(A_d) \\ P(x|B) &= P(x|B \& B_a) \times P(B_a) + \\ &\quad P(x|B \& B_d) \times P(B_d) \end{aligned} \quad (5)$$

式中: x 表示同报文延迟值; A 和 B 代表事件: 报文出现在 A 或者 B 信道上, 而 $A_a B_d$ 代表事件 A 信道先于 B 信道得到调度; $A_d B_a$ 代表事件 B 信道先于 A 信道得到调度.

在式(5)中, $P(A)$ 和 $P(B)$ 与测量器的实际采集点和采集任务有关. 例如在 1:1 的单向随机均衡线路, 被测报文在概率上以等概率分布出现在 A 和 B 两条信道上; 而如果 A 、 B 分别为一对信道的两个方向, 则对于测量单向报文数据而言, 被测报文只可能在某一条信道上出现.

$P(A_a)$ 、 $P(A_d)$ 、 $P(B_a)$ 和 $P(B_d)$ 表示两信道得到调度的概率. 满足且 $P(A_a) = P(B_d)$, $P(A_a) = P(B_d)P(A_a) + P(A_d) = 1$. 在通用硬件架构下, 这与采集任务执行时的各信道的背景流量强度相关. 流量强度大的信道引起的网卡接收中断的频繁度相对高, 从而引起调度概率值的不同.

$P(x|A \& A_a)$ 、 $P(x|A \& A_d)$ 、 $P(x|B \& B_a)$ 和 $P(x|B \& B_d)$ 则与测量器的具体实现相关. 如在上述试验中所得的被测系统体现出稳定的概率特征.

2.2 误差的不确定度表示和传递误差合成

在实际中, $P(x)$ 不能直接用于误差分析和数据处理, 而是需要根据误差理论将其转换成标准均方差 σ 以及最终转化为误差不确定度 Δ .

由于式(5)为加乘关系式, 直接计算 $P(x)$ 的标准方差 σ 非常繁杂. 但由于对于单个报文的到达计算而言, 式(5)中仅 $P(x|A \& A_a)$ 、 $P(x|A \& A_d)$ 、 $P(x|B \& B_a)$ 和 $P(x|B \& B_d)$ 有概率意义, 而其他项可看成非概率的可变项. 可简化其余项为常数项, 转换为 σ 方程为

$$\begin{aligned} \sigma &= \sqrt{a_1^2 P(A_a) + a_2^2 P(A_d) + b_1^2 P(B_a) + b_2^2 P(B_d)} \\ a_1 &= P(A) \times \sigma(x|A \& A_a) \\ a_2 &= P(A) \times \sigma(x|A \& A_d) \\ b_1 &= P(B) \times \sigma(x|B \& B_a) \\ b_2 &= P(B) \times \sigma(x|B \& B_d) \end{aligned} \quad (6)$$

式(6)中 a_1 、 a_2 、 b_1 、 b_2 对于特定的采集任务、测量器接入信道和测量器特性而言均可事先通过计算所得. 只有 $P(A_a)$ 、 $P(A_d)$ 、 $P(B_a)$ 和 $P(B_d)$ 需要在测量期间通过各个信道的背景流量情况计算获得. 通过 σ 计算得出不确定 Δ , 当按照误差处理惯例 $\Delta = 3\sigma$ 是可以在绝大多数情况下保证 $P(-\Delta < x < \Delta) \geq 99.5\%$.

$$\sigma_y = \sqrt{\left(\frac{\partial F}{\partial x_1}\right)^2 \sigma_1^2 + \left(\frac{\partial F}{\partial x_2}\right)^2 \sigma_2^2 + \left(\frac{\partial F}{\partial x_3}\right)^2 \sigma_3^2 + \dots} \quad (7)$$

在协调测量中, 来自不同测量器的数据具有不同 σ . 类似的, 来自不同测量器的时间值 x_1 、 x_2 、 x_3 等合成的间接计算所得值 $y = F(x_1, x_2, x_3, \dots)$ 可以通过误差合成公式得到 y 的标准均方差 (见式(7)^[8]).

和相应的 Δ_y .

由得到的 Δ , 可计算相对误差精度 Δ_y/y 并按照误差测量学误差精度等级来定量判断协同结果是否有意义.

3 结 论

本文分析单点双信道被动测量器的误差来源和成因, 设计了误差分析试验获得被测系统的误差统计特征. 在所得试验结果和对影响测量器实际误差的主要因素分析的基础上, 本文提出了基于概率的单点双信道时间误差分布模型. 利用该模型, 可以在协同测量中使用误差合成手段对间接得出的测量结果的误差进行定量分析, 从而判断测量结果的精度和可行性.

参考文献:

- [1] IP Performance Metrics (IPPM) [EB/OL]. 2004[2005-06-27]. <http://www.ietf.org/html.charters/ippnr charter.html>.
- [2] Packet Sampling (psamp) [EB/OL]. 2005[2005-08-02]

<http://www.ietf.org/html.charters/psamp charter.html>.

- [3] IP Flow Information Export (ipfix) [EB/OL]. 2005 [2005-06-27]. <http://www.ietf.org/html.charters/ipfix charter.html>.
- [4] Cooperative Association for Internet Data Analysis (CAIDA) [EB/OL]. 2004[2005-08-07]. <http://www.caida.org>.
- [5] DEETHS D, BRUNETTE G. [S. l.]: Using NTP to control and synchronize system clocks [M]. Sun Blue Prints OnLine. 2001.
- [6] Computer clock modelling and error analysis. Official NTP documentation [M/OL]. 2005 [2005-08-01]. <http://www.eecis.udel.edu/~mills/ntp/html/index.html>.
- [7] Luca Deri, NETikos, SPA. Improving passive packet capture: beyond device polling [EB/OL]. 2003 [2005-08-01]. <http://www.ntop.org>.
- [8] 杨惠连, 张 涛. 误差理论与数据处理 [M]. 天津: 天津大学出版社, 1992.

A clock error model for single point two channel linked network probe system

GONG Jian, XU Jia-ling

- (1. Jiangsu Province Key Laboratory of Computer Network Technology, Nanjing 210096, China;
2. The East China Regional Network Center of CERNET, Nanjing 210096, China;
3. Computer Science and Engineering Dept., Southeast University, Nanjing 210096, China)

Abstract: Network measurement technology has been one of the hot topics of nowadays Internet technology research. Clock error is the main and direct source of errors in network measurement. Traditionally, the research works on network measurement merely focusing on how to rectify the result and improve clock precision, but neglect the inevitable clock error's effect on measurement result. While this situation is inadaptable for quantified description of time error in distributed cooperative measurement, this paper, integrating the error and precision analysis theory and taking into consideration the features of network passive measurement, represents a clock error model for single point two-channel linked network probe system, and brings in uncertainty to quantificational description of precision and rationality of measurement result.

Key words: computer network; network measure system; clock precision and error analysis; statistics model