

# 如何进行流量的度量

## 问题背景

我们主要是做业务体验模拟的，帮助用户组件网络，给客户提供服务的时候，往往需要现在实验室，根据客户的环境情况、网络需求搭建一个模拟环境，在实验室预先验证新业务的用户感知，但这面临一个难题，现网无新业务平台，而实验室没有现网的背景流量。通常做法是在实验室模拟网络上的背景流，然后叠加新业务，观察用户的体验。

这里面除了设备的问题，我们还有专门的测试仪器模拟用户流量，但是**我们却没有一个很好的模型，或者说有说服力的度量指标去证明给用户，实验室里的流量较好地模拟了真实的线网流量**，具体的来说，模拟的有多好。

## 目前我们是怎么做的

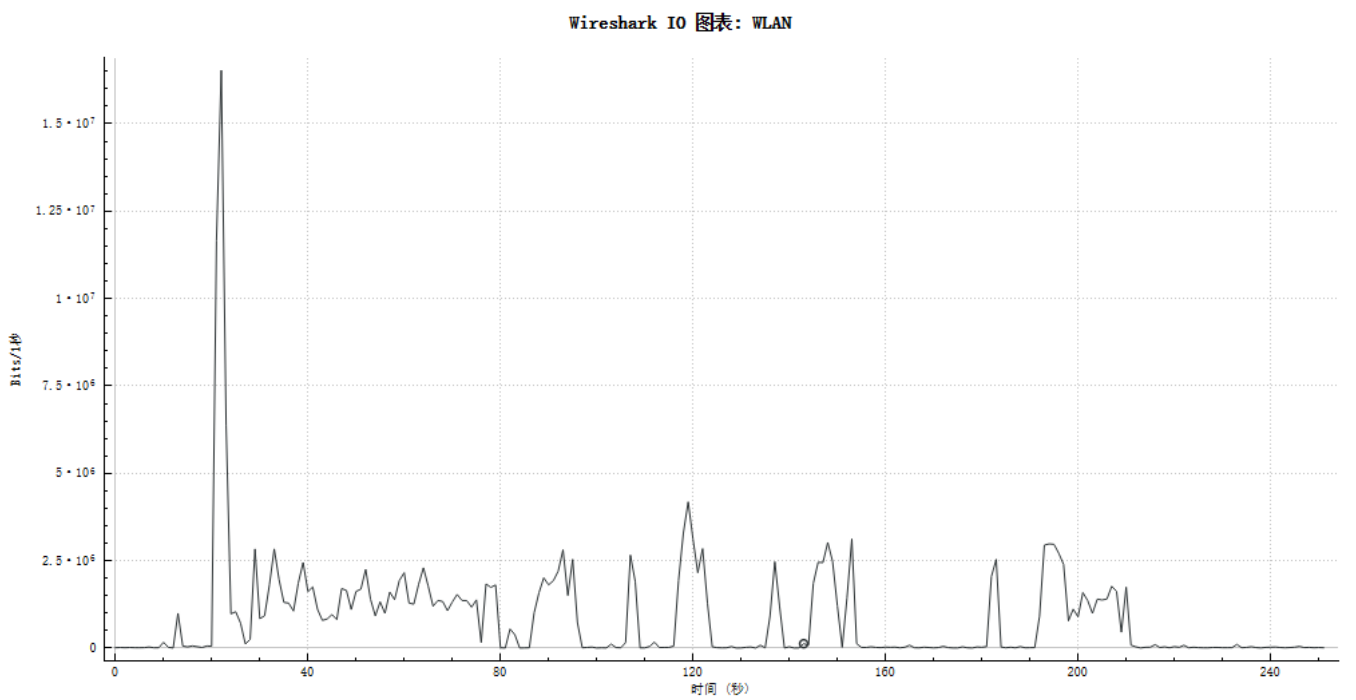
在设备上随机叠加几段业务流，用流量仪设置最大值和最低值用以弥补峰值和谷值的偏差，保证平均值与用户真实网络的是接近的。

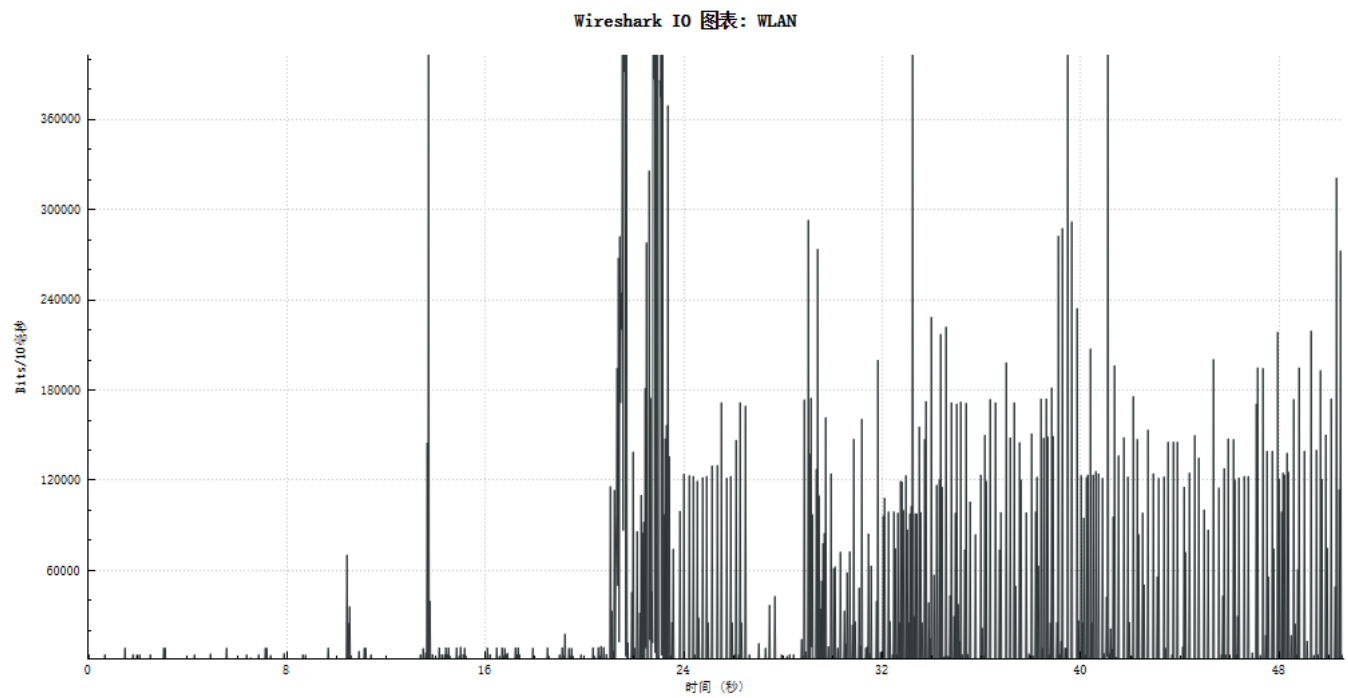
## 数据样本

数据在网络设备上用Waresark软件抓取流量后获得，

- X：表示时间
- Y：表示一定时间间隔通过端口的比特数（bits/时间间隔）我们可以选择不同的时间间隔来分析，比如 **1ms、10ms、100ms、1s、10s、1min、10min。**

下图的图一是每隔1秒进行统计，而图二是每隔10毫秒进行统计。





流量数据可以导出成csv文件进行分析。

## 问题描述

1. **采样问题**：在真实网络环境中采样一段流量数据，然后能在实验室中较好的还原
2. **度量问题**：度量两段流量的相似性
3. **预测问题**：能预测流量固然好，但目前先主要解决一个度量的问题