物联网流量特征

一、数据的来源和网络的结构

数据来源：建立了一个局部的物联网，通过网络包的收集，来得到的数据。

网络建立的结构由图1所示。

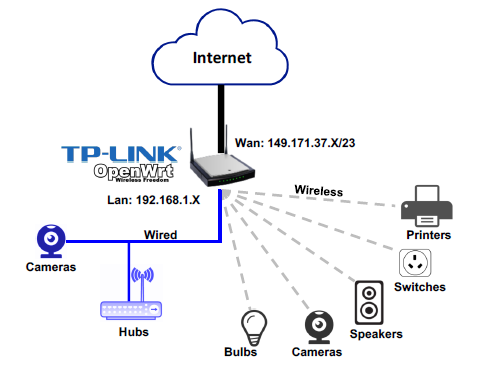
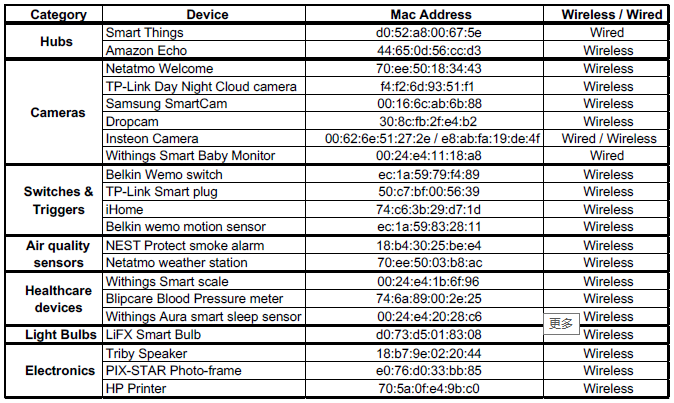


图1 物联网的结构

网络通过路由器来访问外部的网络，在路由器上安装OpenWrt的程序包tcpdump (4.5.1-4)来收集通过该路由器的数据。

物联网连接了21个物联网设备，如图2所示，并且除了以上的IoT设备连接到网络，还有手机、笔记本、安卓平板等非物联网设备。

图2 使用的物联网设备

二、数据分析

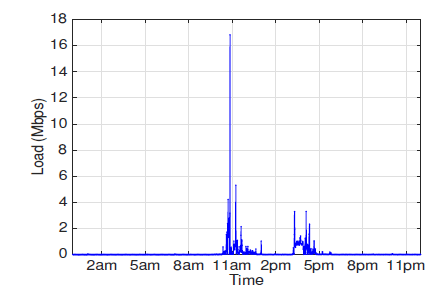


图3 物联网设备和非物联网设备的负载之和

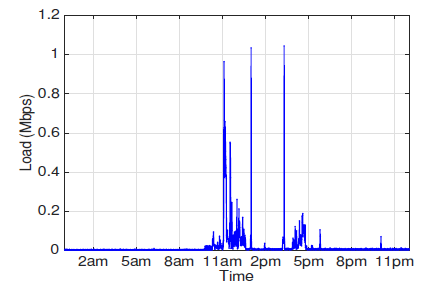


图4 物联网设备负载总和

由图3和图4可知，总体的网络负载峰值在 17 mbps 左右,而平均负载为 400 kbps。

如果我们考虑到只有物联网设备的负载, 峰值负载为 1 mbps 和平均值负载为66 kbps。

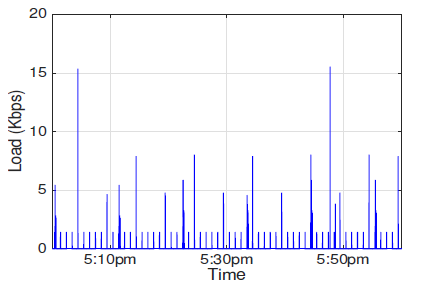


图5 LiFX light bulb

图5代表了LiFX light bulb这个物联网设备的负载随时间的变化情况，可以看出物联网设备生成流量有活动和睡眠期的概念 。

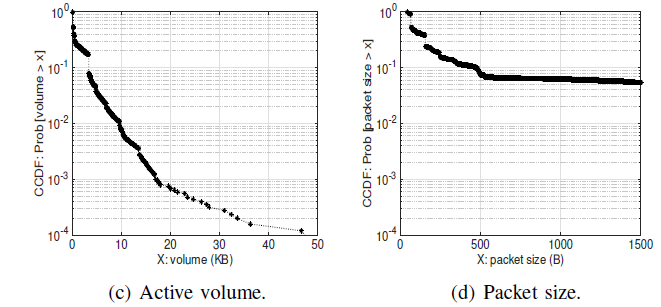


图6 物联网在活动期传送的数据大小和包的大小

由图6可以得出，其中c表明超过75% 的物联网会话传输更少小于 1 kb, 并且只有少于1% 的会话交换超过 10 kb, 这表明大多数物联网设备产生一个小的burst。d表明大约只有10%的包超过500 Bytes。

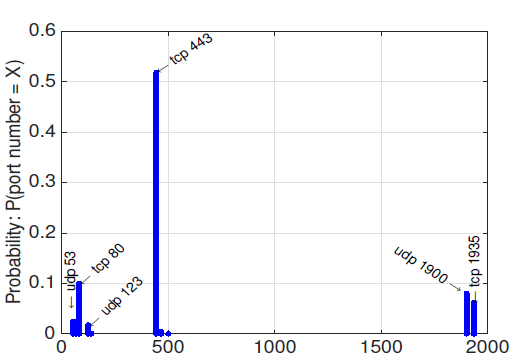


图7 物联网设备选取端口号的概率分布

从图7可以看出，https (即 tcp 端口 443) 是物联网设备主要使用的协议。然而, 约45% 的物联网流量(按数据包数量) 不会通过 https 发送到服务器，这给这些用户带来了严重的安全问题。

物联网设备第二个最主要的应用层协议是 http (即 tcp)), 占物联网流量的11%。

在许多物联网设备中，每个物联网设备平均每天访问的服务器（不包括DNS和NTP）小于10个。而非物联网设备的服务器数量远远高于物联网设备。（比如数据中笔记本电脑会访问100个服务器）

物联网设备仅对一定数量的域名启动DNS查询（主要是其供应商的域名或服务提供商）并以一致的方式重复查询方式。笔记本电脑等非物联网设备会在几个小时的过程中寻找超过300个域名。因此，我们认为，独特域名的数量和DNS查询的频率是物联网设备重要的属性。

笔记本电脑（固定）和智能手机（移动）的流量分析

1. 数据来源

数据主要由两部分组成：WLAN接入点（AP）日志和Netflow记录（），总大小> 30TB。

这些WLAN AP日志是从在大学校园里138栋建筑物中的1760个AP收集的，时间超过479天。它包含超过555M条记录，每条记录包括设备的MAC和分配的IP地址，关联的AP和时间戳。

在2012年4月，对同一网络收集了超过760亿条NetFlow痕迹记录，超过25天。如果一段连续的包具有相同的分组传输协议，源/目标IP和端口号，将被认为是同一个流，由网关路由器收集识别。

笔记本电脑和智能手机的区分：一个设备可以基于MAC地址的前3个字节来识别制造商（带有OUI）。 大多数厂商生产一种类型的设备（只生产笔记本电脑或手机），但有些产品两者（例如，Apple）。

1. 数据分析

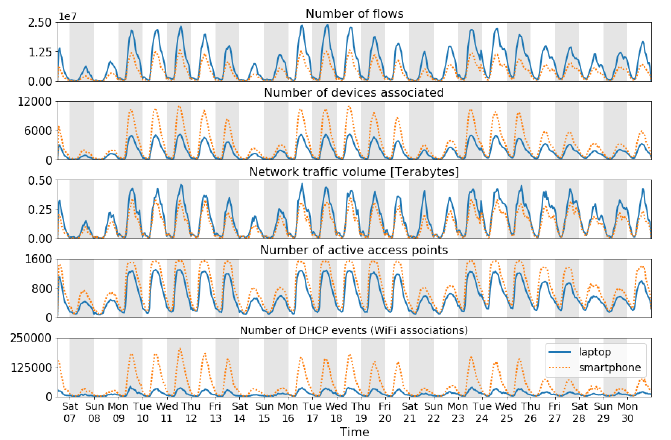


图8 笔记本和智能手机的流量数据

从图8可以看出，尽管智能手机的用户数更多一些，但是笔记本的流量明显的更高一些。同时这些设备的活动有着以天和星期为单位的周期性。

（一）：Flow-leve 的特征（分为智能手机和笔记本电脑两个类别设备所产生的流量数据）：

1) 流的大小是指单个流内所有数据包的大小的总和。手机中单个流的平均大小（2070 bytes）比笔记本电脑的流（822 bytes）大2倍多。手机中流大小的中值（678 bytes）比笔记本的流（142bytes）大4倍多。

2) 在手机中单个数据包的平均大小为（212 bytes）,笔记本电脑中单个数据包的平均大小为(144 bytes)。

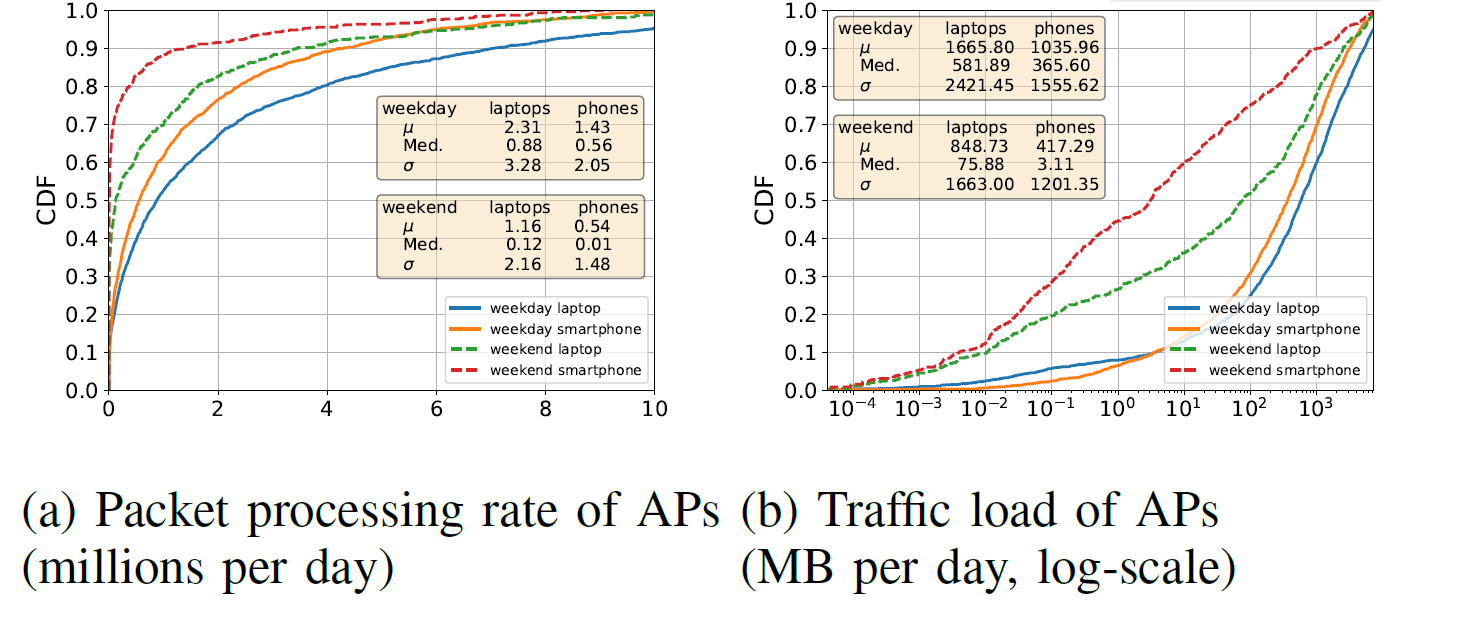
3）手机的单一流包含的数据包个数的均值为7.06，中值为5；笔记本电脑单一流包含数据包的个数的均值为3.64，中值为2。

4）流的运行时间是指流的结束时间减去开始时间。手机的流的运行时间的均值是1868ms，中值是128ms；而笔记本电脑流的运行时间的均值是1639ms，中值是64ms。

5）手机流的分组到达间隔时间的中值为4ms；笔记本电脑的中值为6ms。但是当有许多的智能手机，使得流的到达率很大的时候，手机流的分组到达间隔时间的均值为143ms，笔记本电脑的中值为78ms。

6）手机的流中，TCP流占了78.5%（总bytes占了84.6%），笔记本电脑中，TCP流占了98.2%（总bytes占了91.6%）。由于手机中的UDP的应用更多。在TCP流中数据包的平均个数方面，笔记本电脑流中包的数目是手机的一半（4.6 vs 8.8）；在TCP包的平均大小方面，手机的包比笔记本电脑大22%。

（二）： 网络AP分析：



从a看出，AP每天处理包的数目方面，笔记本电脑的包数目大约是手机的1.6倍。

从b看出，AP工作日每天处理的笔记本电脑的数据小于5GB的概率为90%，处理手机数据小于3GB的数据为90%。