标题：构建WiFi环境下Web用户体验与网络性能的关系模型

引言

研究背景

随着智能手机和平板电脑等智能设备的普及，越来越多的设备只能通过无线保真技术（Wireless Fidelity，以下简称Wi-Fi）或蜂窝网络技术等无线方式接入互联网。同时，因为无线上网的便捷性，越来越多的用户也选择采用无线方式而非传统的有线方式上网。据统计，在2013年整个互联网流量的55%是通过Wi-Fi接入网络的。（I Cisco. Cisco visual networking index: Forecast and methodology, 2013–2018. CISCO White paper, pages 2013–2018, 2013.）不仅如此，支持Wi-Fi的设备从2012年到2017年，预计从50亿台增加至200亿台。（ABI Research. https://www.abiresearch.co m/press/over-5-billion-wireless-con nectivity-chips-will-sh.）这些都表明Wi-Fi已经成为当今互联网最主要的网络接入方式。

相对于传统的有线网络，Wi-Fi网络的环境更加复杂，其服务受到更多因素的影响，服务的稳定性也更加难以保证。在Wi-Fi普及的潮流下，为了给人们提供更方便的上网服务，很多商家都在公共场所布置了大量的无线接入点（Access Point，简称为AP），包括教室、商场、餐厅等。如今的家庭中也会有一个甚至多只多个无线接入点。在给用户提供了方便的上网接入同时，这些无线接入点的相互干扰，可能会影响Wi-Fi的服务性能，给用户带来难以预测的影响。然而由于影响无线网络的因素众多，即使出现上网的性能不能让人满意的问题，用户也难以找到其中的原因。

在互联网众多的服务中，基于HTTP的Web网页浏览服务着数以十亿计的用户，是互联网中最重要也是最基础的应用。（[6] Wang, Xiao Sophia, et al. "Demystifying Page Load Performance with WProf." NSDI. 2013.）网页浏览对于网络性能比较敏感，当网络性能不好时，经常出现网页加载缓慢甚至网页无法加载的问题。这些问题会严重影响浏览网页的用户的体验。用户体验与用户对一个服务提供商的忠诚度至关重要，直接影响着服务提供商可能获得的利润。例如一个用户进行网购时他发现网站无法及时显示他查询的物品或者他进入购物车后迟迟没有反应，他很可能更换到另一个网购网站完成购物，这对于最初的购物网站就是一个损失。统计显示平均增加500毫秒的延迟会导致Bing搜索的利润减少1.2%（S. Souders. Velocity and the bottom line. http://radar.oreilly.com/2009/07/velocity-making-your-site-fast.html, July 2009.），平均增加400毫秒的延迟会导致Google搜索的利润减少0.74%（J. Brutlag. Speed matters for Google Web search. http://services.google.com/fh/files/blogs/google\_delayexp.pdf, June 2009.）。由此可见，Web网页浏览的用户体验关系重大。

虽然Web用户体验非常重要，但是想要简单地获得反应Web用户体验的评价标准却非常困难。与用户相关的信息最主要记录在客户端上，而用户与设备之间的交互更是在客户端才能直接获得。如果希望在客户端上抓取并记录用户的信息用于分析，必须提前让用户安装相应软件，这给大规模部署带来了困难。这些困难导致网络运营商和服务提供商难以有针对性地提升用户体验。

综上所述，Wi-Fi环境下的Web浏览是当今互联网的主要服务，而其用户体验有着巨大的价值和意义。然而，目前缺乏一种有效的在非客户端上测量WiFi环境下Web用户体验的方式。另外，Wi-Fi的多个因素均可能影响用户体验，不同因素对Web用户体验的影响程度也有待研究。本文拟采用决策树和朴素贝叶斯分类基于无线接收点采集的信息构建Wi-Fi参数与用户体验的关系模型，在这个模型的基础上指出对用户体验影响关系较大的Wi-Fi参数，

并给出提升用户体验的指导意见。

相关工作

无线接入点在家庭网络中扮演着一个重要的网络入口的角色，在上面有可能获得更加详细的关于用户设备及用户网页浏览情况的信息，而这些信息是在NAT之后就难以分辨的。基于此，BISmark团队在OpenWRT路由器上部署了Mirage，用于更加精确地监测在NAT之后的家庭网络的性能变化（Measuring and Mitigating Web Performance Bottlenecks in Broadband Access Networks）。BISmark团队在全球超过30个国家部署了测量路由器。他们的研究结果显示接入延迟（Last-mile latency）在整个互联网的延迟中占有很大比例，而接入延迟增加十几毫秒就可能导致网页加载时间增加几百毫秒。他们的实验结果同时显示通过对DNS和TCP连接进行缓存可以有效介绍网页加载时间。

传统研究Web用户体验的方法主要依赖于在客户端安装软件记录用户的信息。WProf（Demystifying Page Load Performance with WProf）是一个植入于浏览器中的插件，目标是分析网页加载时间的不同组成部分。通过监测浏览器的行为，WProf绘制了访问一个网页时网页对象加载、HTML解析、JavaScript与CSS评估和DOM解析的依赖关系图，并且绘制了相应时间的关键路径。关键路径上的节点是构成网页加载时间的不可缺少的部分。研究结果显示网页中的计算部分占据了关键路径时间的大约35%，同时对JavaScript的异步调用以为会阻塞HTML的解析也会显著增加网页的加载时间。

Athula Balachandran（Developing a Predictive Model of Quality of Experience for Internet Video）首次引入机器学习的方法分析网络性能参数和观看视频的用户体验的关系。因为网络参数之间的关联性很强，传统的分析单个网络参数与用户体验的关系难以找到根本的影响因素。他们考虑了视频类型、视频缓冲比、视频缓冲次数、视频平均传输编码率等因素，与用户体验指标——用户观看视频的持续时间进行关联，利用决策树构建了关系模型，并且提出了普遍的用于预测用户体验的框架。

也有工作研究了其他网络类型中网络性能参数与Web用户体验的关系。AT&T团队构建了蜂窝网络中网络参数与Web用户体验的模型。他们在基站上分别采集网络和用户浏览两方面的信息。前者包括基站信号强度、服务用户个数、上下行带宽和不同频率间的切换次数的因素。后者包括网页请求的URL和用户设备的MAC地址等。利用机器学习方法，他们根据信息计算出用户浏览网页的长度、放弃网页的次数和网页未完全加载百分比作为衡量用户体验的标准。在此基础上他们用机器学习的方法构建网络性能参数与用户体验标准的关系。最终，他们发现用户设备在不同频率间的切换对于用户体验有最大的影响，因为频率间的切换会导致网络中断长达几秒钟的时间，接下来影响较大的因素是信噪比。与预期相反，上下行带宽对用户体验影响不大，可能原因是网页浏览对带宽的需求较低。

主要工作

本文的工作主要分为以下几个部分：

1. 实现基于OpenWRT路由器的测量工具：

在路由器上抓取网络相关参数

为了构建网络性能与用户体验的关系，首先需要获得在用户使用无线网络时的路由器参数，包括时间戳、信道占空比、上下行带宽、物理层上下行速率、MAC层数据包重传比例、路由器当前连接设备数和每个设备的信号强度等。获得的数据传输到服务器上用于进一步分析。

在路由器上抓取用户访问信息

目前的Web应用基本上基于HTTP完成。为了了解用户的网页浏览状态，需要在路由器上抓取HTTP数据包记录。具体包括HTTP中GET类型数据包的GET、Host、User-Agent和Referer域，以及设备的MAC地址和时间戳。获得的数据同样传输到服务器上用于进一步分析。

1. 实现用户体验的分析工具

构建网站的HTTP请求的下载链

路由器上抓取的HTTP的GET数据包不仅来源于用户的网页浏览，也可能来源于其他Web应用。为了研究用户的网页浏览体验，首先要过滤掉其他来源的数据包。因为在路由器上抓取的记录是以数据包为单位的，在过滤掉无关数据之后，还需要在服务器上利用抓取的数据包的关系，恢复出用户对网站的完整访问流程。

区分用户请求与网页嵌入对象的自动请求

网页浏览的HTTP的GET请求来源分为两种。一种来自于用户点击一个超链接或者在浏览器中输入一个网址导致的主动请求。另一种来源于浏览器解析网页时，针对嵌入的对象如图片、JavaScript和CSS等自动发出的请求。前者与用户网页浏览时的投入程度关联紧密即与用户体验直接相关。后者主要用于呈现网页的内容，并不直接反应用户体验。因此，需要在请求中对两者进行区分。

构建网络参数与用户体验的关系模型

在区分出用户请求之后，将其作为衡量用户体验的标准。把用户请求与对应的Wi-Fi网络参数匹配，分析网络参数对其不同的影响程度。在此基础上挑选出具有主要影响力的网络参数用于构建关系模型，并且基于模型提出提高用户体验的建议。

1. 为了能够将带有测量工具的路由器分发开展实际测量，首先需要完成测量数据与服务器间的传输工作。然后，需要控制测量工具的计算开销。在实际部署测量之后，需要根据收集到的数据进行分析，进一步完善测量程序。

论文结构

本文第一章引言介绍了相应的研究背景、已有的工作和文章的主要工作内容。

本文第二章主要介绍不同的测量Web用户体验的方法。

本文第三章主要介绍系统的架构及实验的部署。

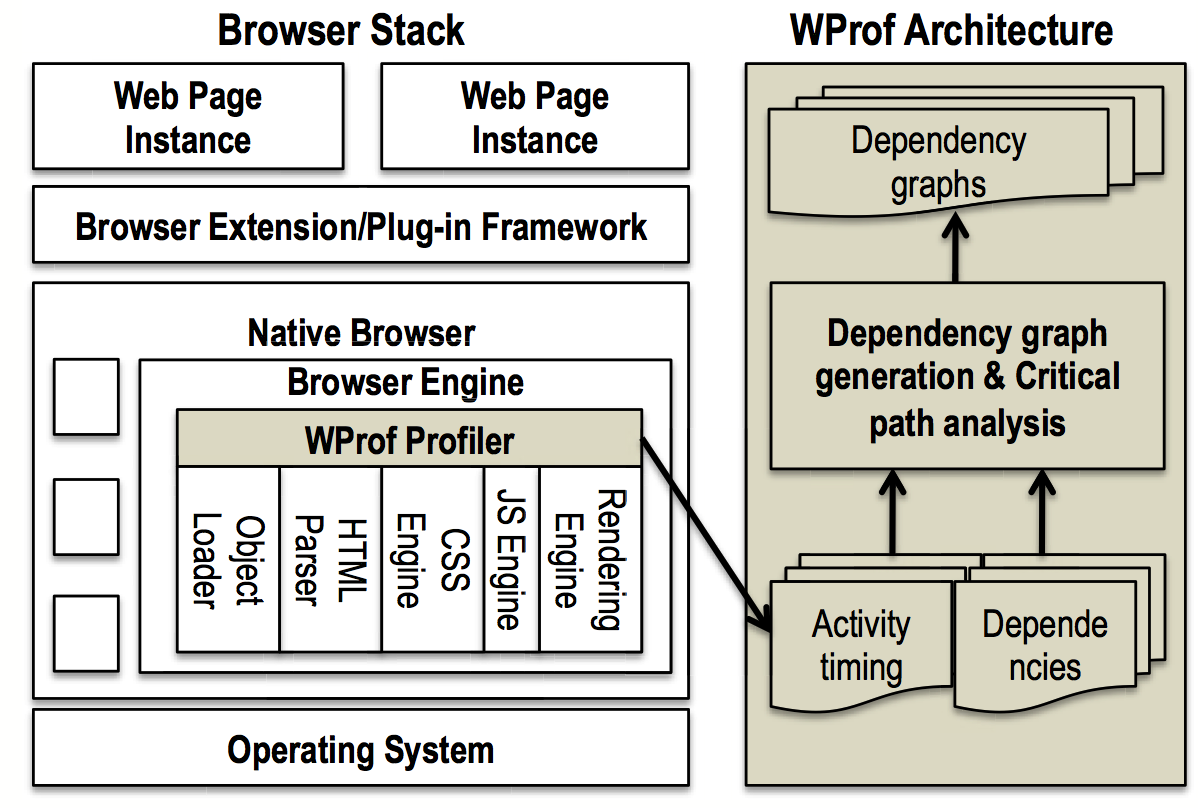
本文第四章主要介绍模型的构建以及数据的分析。

本文第五章对文章进行总结，探讨实验中遇到的问题以及未来的后续工作。

第2章Web用户体验的测量

本文的主要目的是构建Wi-Fi网络性能与Web用户体验的关系模型，准确有效地测量Web用户体验是构建模型的前提与基础。测量的方式也决定了需要收集的数据和实验框架的构建。

本章节首先根据测量方式的分类具体介绍了几种不同的测量方式，然后详细介绍了实际采用的一种适于大规模部署的基于路由器被动测量的方法。

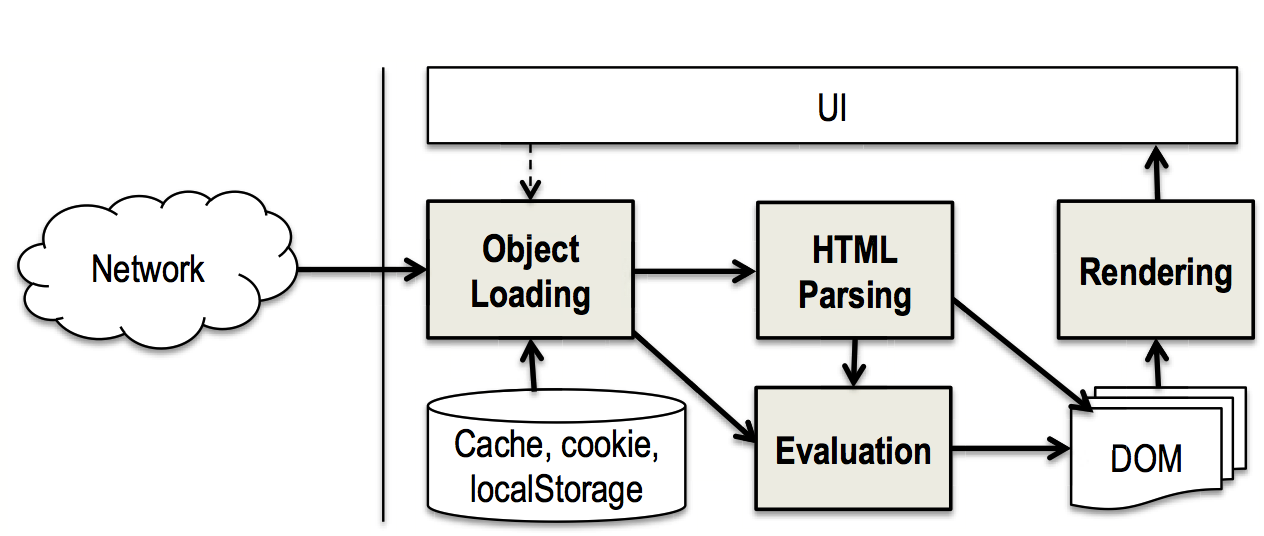


主动实验者实验

一些关于用户体验的实验邀请实验者到实验室中完成规定的任务，通过实验后的问卷调查了解用户体验。DF Galletta请196名学生在实验室环境下完成对三个网站的搜索任务。三个网站保存在光盘上，用以严格控制访问时间。在用户完成访问不同的访问任务时，电脑加入人工设置的延迟时间模拟不同网络性能。在用户完成实验后，他们被要求填写问卷评估自己完成任务的目的性和感受。主动实验者实验的优点在于调查问卷获得的用户反馈最真实直接地反应用户体验，是最准确的标准。然而这种实验方式缺点也非常明显，实验需要在实验室环境中才能完成，因此地点受到限制，导致被试者的普遍性受到限制。除此之外，每一次的实验代价非常大，持续时间却不能太长，不可能进行大规模实验。

客户端上的测量

目前常用的一种测量方法是在用户的设备即客户端上安装专门的程序捕捉用户信息或者用户与设备直接的交互情况以此来推测用户体验。Xiao Sopiha Wang将WProf，一个网页加载分析插件，植入于不同种类的浏览器的引擎中。该插件可以将浏览器加载网页的每一个事件记录下来并根据发生时间和相互影响构建它们的依赖关系。在客户端上进行测量的优势是可以获得客户端上独有的数据，在该种情形下比如网页的渲染时间和浏览器对JavaScript的计算时间。这些因素构成了网页加载时间的重要部分，直接影响着用户体验，而且这些数据只能在客户端被获取。在客户端进行测量的最大的缺陷是需要让用户额外安装测量软件，这种不便使得该种测量方法难以被大规模部署，也就难以适用于广泛地获得大量数据进行实验。另外一个缺点是在用户安装测量软件后，用户对隐私被泄露的担心可能改变他们使用网络的习惯，不能最真实地反应实际用户使用情况。



服务器端的测量

中间节点的测量

被动测量

基于中间节点的测量

鉴于本文探讨的是构建Wi-Fi环境下网络性能与Web用户体验的关系，我们将中间节点限制在普通类型的家用路由器上。路由器作为室内无线上网的主要接口，通常是室内无线上网所有流量经过的通道。同时，路由器作为接入互联网的最后一跳，它与用户设备直接相连，在路由器上收集数据使得我们能够将不同的记录针对不同设备或用户区分开，给后续研究奠定基础。从路由器之后这些数据包被汇聚在一起，弱化了每个用户独有的特征。因此，路由器是部署Wi-Fi性能测量工具的最佳选择。在路由器上，采集到的Wi-Fi参数主要分为两类。一类是路由器的总的工作状态记录，例如路由器累计的工作时间、累计传输数据的时间、累计接受数据的时间等。另一类数据是与单个设备相关的，包括当前连接的设备个数、设备MAC地址、每个设备传输的数据量、每个设备接受的数据量和每个设备的信号强度等。

路由器虽然是理想的测量Wi-Fi性能参数的地点，在测量用户体验时却不够直接。因为用户体验是用户层面的一种衡量标准，更加宏观，而路由器上只能获得包级别的底层数据。这就需要一个机制能够从底层的数据包记录中恢复用户体验的相关数据。

我们选择的衡量Web用户体验的指标是会话长度（Session Length）（引用！！！）。会话长度指用户进入一个网站后，从该网站及相关链接中手动打开的网页个数。会话长度表示了用户的投入程度。当网络性能良好时，网页打开的过程非常流畅，用户体验良好，因此会使用网页直到达到最初的目标。当网络性能不好时，网页的加载过程可能非常缓慢，有可能出现网页部分加载失败甚至整个网页无法加载的情况，这时用户的浏览体验不好就会关闭网页，导致会话长度很短。因此会话长度可以有效衡量Web用户体验。

接下来需要利用包级别的数据推测会话长度。用户手动打开网页反映在数据包记录中就是一个针对网页地址的HTTP的GET请求。然而路由器上记录的HTTP的GET请求并不仅仅来自于用户手动点击网页的访问。当浏览器解析网页的HTML文件时，会对其中的嵌入对象比如图片、JavaScript和CSS发出HTTP的GET请求，下载内容完成网页的加载，这些请求占据了GET请求的一大部分。现在主要问题转变为如何区分用户主动点击发出的请求和网页嵌入对象自动发出的请求。一种传统做法是（引用！！！）利用请求的间隔时间（idle time）。这种方法做出了以下假设，因为对嵌入对象的请求是在浏览器解析网页时自动发出的，时间间隔非常短，请求一般都在网页加载时发出。相反，用户的点击表示打开新的页面，一般在浏览原有网页之后，相隔时间较长。因此，通过对比两次相邻链接之间的时间差，如果大于预定阈值就可以认为这个请求是用户点击产生的，反之就是浏览器发出的。

然而，如今网页的复杂度逐渐增加，需要动态加载的内容越来越多，加载时间的跨度也逐渐增加，利用一个固定的阈值已经难以区分网页嵌入对象自动发出的请求与用户点击的请求。本文采用了一种更加有效的基于文本分类的用户点击识别方法。观察从一个网站中发出的所有请求，可以发现请求的URL有明显的区分。以新浪网为例，www.sina.com.cn和sports.sina.com.cn很可能是用户点击的网站链接。而对sax.sina.com.cn或d1.sina.com.cn的请求很大可能性来源于对嵌入对象的请求。基于上述特征，文本分类的方法可以应用于区分用户点击去嵌入对象。我们分解URL的中间部分以及文件类型作为特征。以下面这个URL为例http://www.cs.tsinghua.edu.cn/publish/cs/index.html，特征集为{cs, tsinghua, edu, cn, publish, html}。分类方法上，我们选用了在文本分类中效果很好的朴素贝叶斯方法。为了减少人工标注训练集的工作量，我们采用了一种自动标注的方法。将从一个网站开始的全部请求按时间排序，取出前10秒的部分。如前所述，因为人在每次手动点击之间的时间是浏览网页，相对较长，我们可以认为10秒这个区间内只有第一次请求为用户点击发出的，其他请求都为浏览器自动发出的。我们可以使用这样的方法自动生成训练集。获得足够的训练集之后，用这些训练集训练朴素贝叶斯模型，然后用训练后的模型对所有请求数据进行处理，获得仅包含用户点击的列表。