关于清华校园无线网络安全性的研究

2013 年 11 月 25 日

中文摘要

清华大学校园网络发达，无线热点覆盖较为全面，为广大师生带来了极大的便利。然而，这种广泛性对网络的安全性提出了要求。本文在尝试从理论上分析校园无线网络的安全性的同时，从实践上给出了截获（监听）校园无线网通信数据的可能性的实验证据。

在截获通信数据的基础上，本文还分析了利用无线网络登录界面的严重编程失误来窃取登录者的校园网账户名和密码的方法。

本文不但分析了无线校园网的安全问题，还给出了两种不同的解决方案，即临时性的解决方案与长期的解决方案，可以为今后校园无线网络的建设工作提供参考。

关键词：无线校园网；网络安全；网络数据截获

ABSTRACT

The wireless network, which covers large areas of the school and bring great convenience to both teachers and students in Tsinghua campus is very advanced. The extensive usage of the network, however, requires good security. This paper tries to discuss the security of campus wireless network in theory, while the evidence of the possibility of capturing the data transferred via the network is also given in practice.

Based on the ability of capturing data, the paper also analyses the method of taking advantage of the serious programming mistakes of the login interface of the campus network to steal campus accounts and passwords.

The paper not only analyzes the security problem of the campus wireless network, but also gives two solutions, one is temporary and the other is permanent. The solution can be a reference to the future construction of Tsinghua campus wireless network.

Keywords: campus wireless network;network security;capture of data transferred via network

目 录

主要符号对照表

OSI	开放式互联 (Open System Interconnect)
Wi-Fi	无线传输系统
AP	无线访问接入点 (WirelessAccessPoint)
WPA	保护无线电脑网络安全系统 (Wi-Fi Protected Access)
SSID	服务集标识 (Service Set Identifier)
HTTP	超文本传输协议 (Hypertext Transfer Protocol)

第 1 章　数据截获的条件与原理

清华校园无线网是基于 SSID 为 Tsinghua 的 Wi-FI 网络的。无线网络的安全性的问题源于，数据是以无线信号的形式在终端（即上网设备）与访问点（AP）之间进行传输。原则上，无线信号是一种电磁波，它与有线通信不同，它不是定向传播的。这就为信号的窃听与截获创造了方便，但是窃听与截获能否成功，是取决于多方面的因素的。

1.1　Wi-Fi 网络的结构与通信机制

OSI 模型指出，计算机网络体系结构划分为以下七层[1]：

- 应用层
- 表示层
- 会话层
- 传输层
- 网络层
- 数据链路层
- 物理层

Wi-Fi 则是工作在数据链路层的一种使得设备能够以无线形式传输数据的技术。顾名思义，Wi-Fi 是一种使得数据能够得以传送的通信协议。无论 Wi-Fi 的通信机制如何，按照 OSI 模型的约定，数据链路层只负责透明的对数据进行可靠的传输[2] 。这里的透明是指数据链路层不必对数据进行解析等工作，只负责传输就可以了，可靠是指保证数据完好地传输到目的而不至发生丢失或者混乱。从这里我们不难看出，OSI 模型并没有对数据链路层对于数据传送的具体实现方法以及其保密性做出要求。

建立在数据链路层上的依次是网络层、传输层、会话层、表示层以及应用层，根据层与层之间的透明性的规定，这些层不能也不必知道数据链路层究竟是以什么方式实现的。这些层与普通的有线网络没有什么不同之处。

由于无线信号的特点，在网络的工作过程中，从 AP 发向某一特定装置的数据信号会被所有连入的设备接收到，但是（在正常状况下）只有该特定装置会

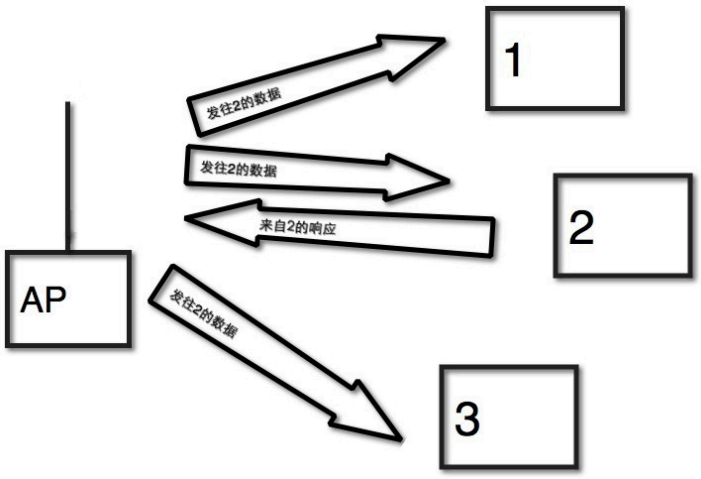

图 1.1　广播的示意图

接受并响应该信号，而其他设备则会忽略这一信号。这一机制被称作“**广播**”，如图 1.1 所示。

1.2　数据截获的原理

正如 1.1节所述，在广播的机制下，无论发往哪一个设备的数据都会被传送到网络内部所有的设备中，这就为数据截获带来了可能。如果我们设法使原本应该被忽略掉的数据没有被忽略掉，而是被我们收集并处理，那么从原理上，我们就实现了对整个无线网络的监听，如图 1.2 所示。但是，事实上我们只是在数据链路层截获数据，而数据链路层只是负责数据完整的的传输，并不负责对数据的解释，因此，我们捕获记录下来的数据需要我们自己来分析。如果 Wi-Fi 传输数据的时候采取了某种加密措施，那么我们获得的数据就是一串无意义的（对于我们来说是不可理解的）数据，而无法解读其内容，从而使得截获失败。

1.3　数据截获的条件

综上所述，我们要想截获无线网络的通信数据，需要满足：

- 数据基于广播机制进行传输；
- 收到的目的地不是本机的数据没有被忽略；
- 数据链路层没有加密。

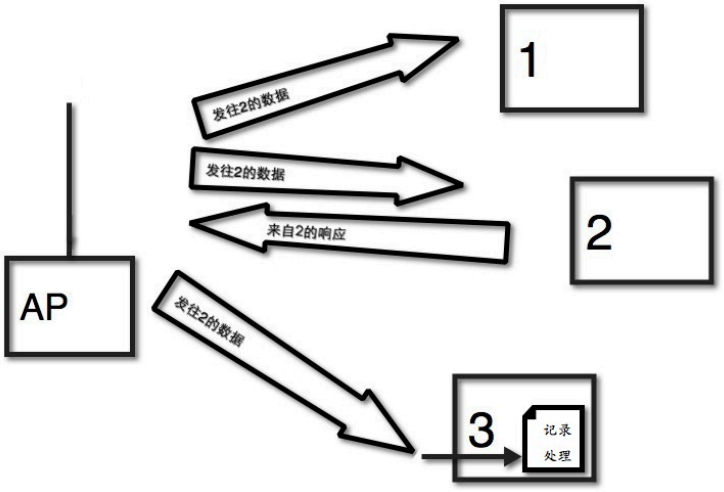

图 1.2　广播环境下数据截获的示意图

在下一章中，我们会具体地分析这些条件是如何在清华无线校园网中得以满足的，并且给出一个实际可行的截获的方法。

第 2 章　无线数据监听的实验

本章将会具体的讨论无线数据监听的可能性和方法。并且以清华大学校园无线网络为具体实例，来演示数据监听（捕获）的操作方法。

2.1　可行性分析

我们来分析在 1.3 节中指出的几个条件能否在清华无线校园网中得到满足。

2.1.1　广播机制

这一点毫无疑问是满足的。在无线信号的发送过程中，显然无法建立起一条从 AP 到终端设备的专用链路①。数据的传送只能（在无线的条件下无法避免的）通过广播来实现。

2.1.2　监听模式

一般情况下，数字链路层会分析所收到的数据包②，并过滤掉目的地不是本机的数据。这实际上使软件系统无法获取所有的数据包。但是，这个困难是十分容易解决的。通过一定指令，可以使网卡进入“**监听模式**”(如图 2.1)。在监听模式下所有收到的数据都会被呈递给网卡之上的层次进行分析处理[3]。因此，在“监听模式”下，我们可以看到网卡接收到的所有数据。

2.1.3　开放网络

清华校园无线网络，即 SSID 为 Tsinghua 的无线网络没有采取任何加密措施（如图 2.2），这时，我们就可以在数据链路层上截获所有的传输数据③，并且在其中挑选我们感兴趣的数据进行进一步的处理。

---

① 然而，在有线网络中，从交换设备到终端设备的物理线路天然的形成了“专用链路”。因此，为了避免广播的潜在不安全因素，需要做的只是使接入设备能够小心的选择数据应该发送的端口，从而使得终端设备根本接收不到不是发往它的数据信号。这一技术称为**交换**，实现了这一技术的接入设备称为交换机。目前，绝大多数有线网络都是用交换机进行组网，集线器已经逐步被淘汰。

② 一般由网卡实现。

③ 这意味着如果数据在其它（更高）的层次上被加密了，我们仍然无法获知其真实内容。

图 2.1　处于监听模式下的网卡

图 2.2　开放的 Tsinghua 网络

2.2　监听实验

2.2.1　实验条件

2.2.1.1　时间、地点

2013 年 11 月 23 日 19:00 至 21:00 人文社科图书馆三层

2.2.1.2　器材

笔记本电脑一台，系统为 Mac OS X 10.9，无线网卡是 Broadcom BCM43xx 1.0 (5.106.98.100.22)。所用软件为 Wireshark[4] Version 1.10.2 (SVN Rev 51934 from /trunk-1.10)。

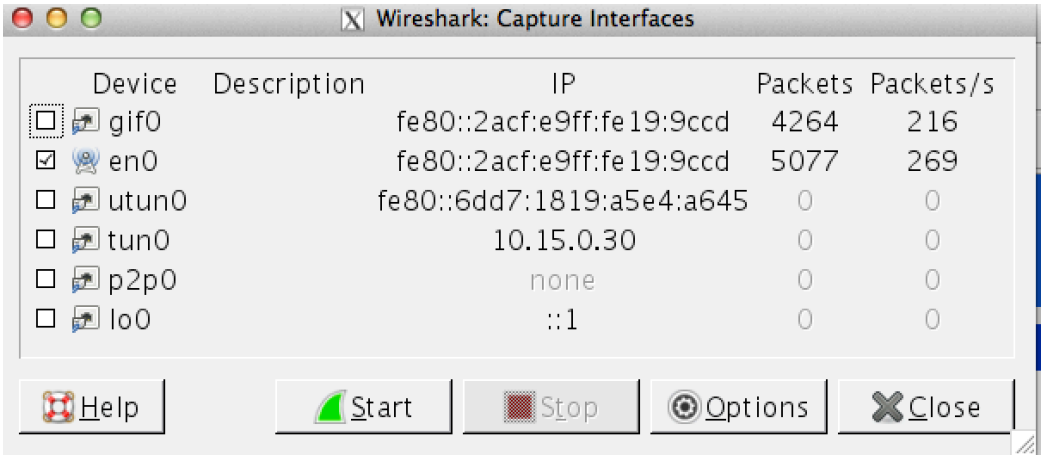

图 2.3　配置网卡监听参数（1）

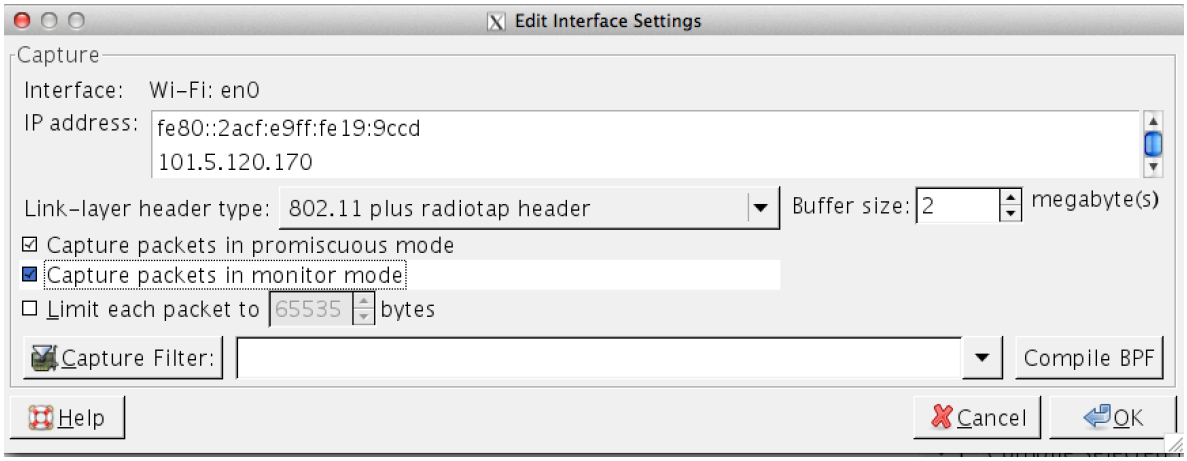

图 2.4　配置网卡监听参数（2）

2.2.2　实验方法

(1) 启动 Wireshark

(2) 在界面左侧选择“Interface List”

(3) 在打开的对话框中选取 Options（如图 2.3）

(4) 在网卡列表中双击选择“en0”

(5) 勾选“Capture packets in monitor mode”，然后点击 OK （如图 2.4）

(6) 点按“Start’”按钮开始监听

(7) 一定时间后点按工具栏中的 Stop 按钮结束监听

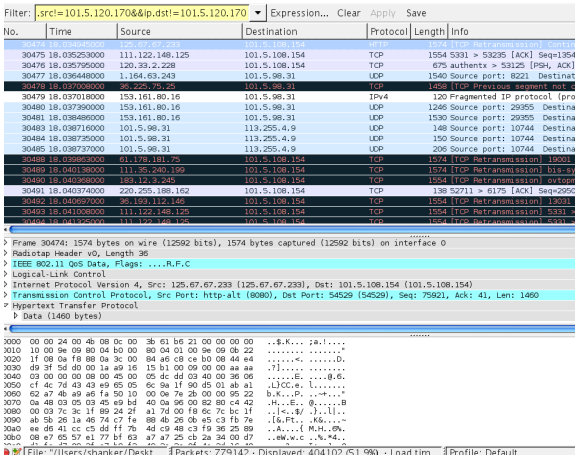

图 2.5　截获的数据

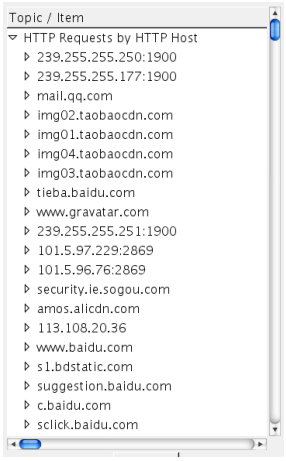

图 2.6　截获数据的分析

2.2.3　实验结果

在实验中截获了近 2G 的数据（如图 2.5），这些数据中有 99.3% 是来自其它设备的，从这些数据中我们可以得到很多信息，包括用户访问的网站，浏览的页面（图 2.6）等等。如果这些页面中含有用户的隐私信息，那么这些隐私信息将会暴露无遗，这是十分危险的。我们会在之后的章节中结合校园网登录页面的严重安全问题进一步分析这些数据。

2.2.4　实验结论

实验结果印证了我们在第 2.1节中的分析，即校园无线网络可以被监听，并且，明文数据占了相当大的比例，因此，从网络的保密性的角度来看，清华校园无线网络是不够安全的。

第 3 章　监听数据的利用——以校园网账号为例

在前两章里，我们重点关注了数据链路层，即数据捕获的原理与方法，当已经捕获到了数据的时候，我们该如何利用这些捕获到的数据？这就涉及到对应用层的应用协议的分析以及利用其弱点的过程。本章将会着重讨论如何利用捕获的数据来找到我们感兴趣的信息。当然，为了具体地说明问题，我们不妨假定，我们感兴趣的信息是同学们校园网的账号和密码。首先分析的是校园无线网登录页面的特点及弱点，之后我们会讨论如何有针对性的利用这些特点来分析筛选我们之前截获的数据。

3.1　HTTP 协议

HTTP **协议**[5]（Hypertext Transfer Protocol）的全称是超文本传输协议。这种协议目前被广泛的应用，我们日常浏览的网页就是通过这种协议传输的，本节将会简要地介绍 HTTP 协议。

3.1.1　请求

HTTP **请求**（Request）是指浏览器向 HTTP 服务器递交的，说明浏览器想要请求的资源的数据。HTTP 请求分为头部和正文两部分。其中，正文在没有附加数据的情况下可以省略。下面是一个简化的 HTTP 请求样本：

```
1  GET /index.php HTTP/1.1
2  Host: localhost
3  Accept-Encoding: gzip, deflate
4  User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS
   X 10_9) AppleWebKit/537.71 (KHTML, like Gecko) Ver-
   sion/7.0 Safari/537.71
5  Accept-Language: zh-cn
6  DNT: 1
7  Connection: keep-alive
```

```
8
```

这里必不可少的是第 1 行和第 8 行，第 1 行中 `GET` 指明了请求方式为 GET，即不向服务器提交附加数据，`/index.php` 是请求的资源，`HTTP/1.1` 指明了 HTTP 协议的版本为 1.1。最后的第 8 行是空行，表示头部的结束。第 2 行至第 7 行是头部指令，指明了请求的附加要求和浏览器的相关信息。

3.1.2 响应

HTTP **响应**（Response）是指 HTTP 服务器对请求的应答，HTTP 响应中包含状态指示（请求成功或请求失败）、附加说明以及所请求的资源等。以下是一个 HTTP 响应的样本：

```
1  HTTP/1.1 200 OK
2  Date: Sun, 24 Nov 2013 09:05:12 GMT
3  Content-Type: text/html;charset=utf-8
4  Expires: Sun, 24 Nov 2013 09:04:18 GMT
5  Cache-Control: private
6
7  <!DOCTYPE html>
8  <html><head>
9  ...
```

响应的第一行说明了 HTTP 协议的版本（1.1）、状态码（200，响应成功）以及对状态的简短表述（OK）。接下来的 2 至 6 行是标头指令，对响应做出了进一步的说明，包括内容格式，数据长度等，这部分是可选的。第七行是空行，作为对标头和内容的分隔。其余的部分就是响应正文。

3.1.3 维持状态

HTTP 是一种无状态协议，即在一次请求与其响应之后，浏览器与服务器的连接会中断[①]。但是在实践中，我们需要知道连续的几次请求是否来自同一

① 目前新的标准可以使连接持续，以便服务器能够多次“推送”数据给浏览器，但是，在一次会话中，请求只能递交一次。

用户[①]。这就需要 HTTP 服务器发给浏览器某种数据作为凭据，并且要求在以后的某段时间内[②]在所有的发往这台服务器的请求中，都附加上这个数据，以便服务器识别。这种数据，就被称为 Cookie。

Cookie 的传送都是在 HTTP 标头中头部指令中以明文方式传送的。

3.2 清华校园无线网关的登录页面

3.2.1 通信协议

当未认证用户接入无线网之后，如果这个用户尝试访问网络[③] 那么他就会被转向登录页面（如图 3.1）。与这个页面有关的数据传输过程[④]都是通过 HTTP 协议进行的。然而，HTTP 数据是明文传输的。从 OSI 模型上来看，HTTP 是在应用层上的。也就是说，我们在应用层上的安全已经丧失了。然而，根据第 2.1.3节的讨论，我们可以看到，在清华无线网中，数据在数字链路层也是没有加密的，这就导致了浏览器与服务器的通信数据暴露在“大庭广众”之下。

3.2.2 通信内容

这些浏览器与认证服务器之间的通信数据中，是否含有我们所需要的数据？这是一个十分关键的问题。这就需要我们对登录页面作出分析。

我们来看登录部分的代码：

http://net.tsinghua.edu.cn/script/login.js

```
1  include('/script/cryptojs.md5.js');
2
3  function do_login() {
4      var uname = $('#uname').val();
5      var pass = $('#pass').val();
6
```

① 比如在需要登录的页面我们需要知道发出请求的用户是否已经登录过。如果没有维持状态的技术，那么这个问题是无法解决的。

② 在技术上我们称之为**有效期**

③ 指浏览网页

④ 例如，页面发送给浏览器，用户名密码发送给服务器进行验证.

图 3.1　校园无线网的登录页面

```
...
18
19      var topost = "username=" + uname + "&password=" +
    CryptoJS.MD5(pass) +
20          "&drop=0&type=1&n=100";
21
22      var res = post('/cgi-bin/do_login', topost);
23
...  //... 略去以下代码
```

当用户点按“连接”按钮时，这里的 `do_login()` 函数就会被调用函数的第 19 行是产生提交的数据的，然后这些数据在第 20 行中被提交到了服务器上进行验证。但是，请注意，生成的提交数据中的密码是经过 `MD5()` 函数进行散列运算的。因此，登录页面并没有把密码的明文传送出去，而是传送了密码的 MD5 散列值，所以，我们仅能从截获的数据中得到散列值而得不到明文密码[①]。但是着并不代表这个登录界面就是安全的。**服务器永远不会知道浏览器中究竟发生了什么**，我们完全可以提交自行构造的数据，在这个构造的数据中直接填写获知的散列值，从而达到通过登录验证，盗用他人上网流量的目的（如图 3.2、3.3、 3.4、 3.5）

① 虽然我们可以得到散列值，但是，由散列值得到明文在数学上是不可能的。

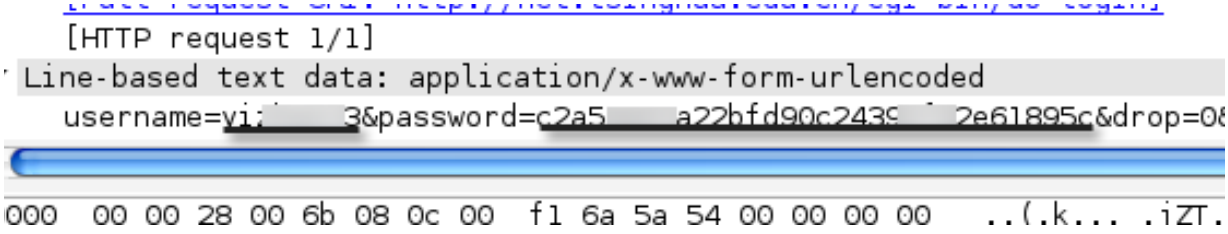

图 3.2　截获的账户名和密码的 md5 散列

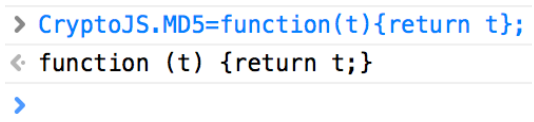

图 3.3　替换 md5 散列函数

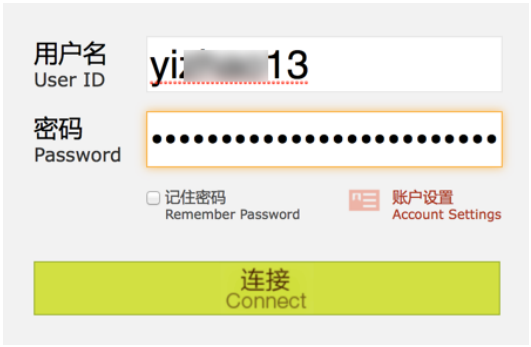

图 3.4　在页面中输入账号名和密码散列值

图 3.5　流量盗用成功

3.2.3 记住的密码

同时，我们还可以发现，登录页面上还有一个功能是记住密码。每次打开这个登录页面的时候，我们之前登录时使用的用户名和密码就会被自动的填写到表单中。当分析这个功能的实现原理的时候，我们会发现其中的严重安全问题。我们先来看页面载入时的代码

http://net.tsinghua.edu.cn/script/login.js

```
112  $(document).ready(function() {
113       var cookie = $.cookie('tunet');
     //获得名称为 tunet 的 cookie
114       if (cookie) {
115            var a = cookie.split('\n', 2);
     //用换行符分割
116            $('#uname').val(a[0]);
117            $('#pass').val(a[1]);
     //将分割的两段分别填入用户名和密码的相应空白
118            $('#cookie')[0].checked = true;
     //勾选“记住密码”复选框
119       }
...
131  })
```

每当页面刚刚载入完成的时候，这段代码便会执行起来，完成将密码填入的操作而 cookie 中的数据则是由登录时执行的程序存储的[①]。

http://net.tsinghua.edu.cn/script/login.js

```
 3  function do_login() {
...
25           if ($('#cookie')[0].checked) {
     //判断“记住密码”复选框的状态
```

① 客户端脚本 Javascript[6] 也可以实现对 cookie 的操控

图 3.6　截获的明文密码

```
26              $.cookie('tunet', uname + '\n' + pass,
    //直接将用户名和密码用换行符连接起来保存到 cookie 里
27                      { expires: 365, path: '/'});
28          } else {
29              $.cookie('tunet', null);
30          }
...
```

我们注意到，在这里，cookie 并没有发挥我们在第 3.1.3节中说明的作用。它仅仅被用来在本地[①] 存储数据而已。且不说明文存储在本地的 cookie 会不会很轻易地被非法获取，就从网络监听的角度来分析，这也是及其**危险**的。按照 cookie 的标准，无论是什么来源的 cookie，都要在“所有的发往这台服务器的请求中”，发送这个数据。这意味着，虽然并没有理会[②]，但是，存储在 cookie 中的数据仍然会以**明文**的方式发送给服务器。于是我们就可以从截获的数据中发现账户名以及**明文**的密码（图 3.6）。

相比于第3.2.2节所叙述的安全问题而言，本节所述的安全问题更加严重。因为前者只是盗用校园网流量而已，而利用后者截获的密码则可以进一步地进行恶意的破坏，如盗用清华学生邮箱、浏览信息门户的隐私信息、恶意退课等。

3.3　从截获的数据中筛选敏感信息

如第 2.2.3节所述，我们在实验的过程中得到了海量的数据，对于这些数据，我们应该进行筛选，选择出我们想要的数据出来。

根据上面的分析，我们需要搜索的数据应当是发往验证服务器的 HTTP 请

① 指浏览器

② 因为当记住密码时，用户名和散列过的密码与没有记住的时候一样，仍然会被发送给服务器进行验证。因此，我们不难肯定这种猜想，即服务器根本没有理会浏览器发送的 cookie

求。这样，我们不难构造一个过滤器[①]：

```
http.host contains "net.tsinghua.edu.cn"
```

只需要将这个过滤器输入到 Wireshark 的过滤器栏中，即可筛选出符合条件的数据。在这里，为了表明校园无线网络安全问题的严重性，我把所有截获的用户名和密码罗列如下：

表 3.1　截获的账号名和密码

账号名	密码
t**13	dan****01
c**11	1MA********NCAI
l***s11	77****8ls
s****c12	41********chi
h****1	1988****jun
ho****kv10	EN****8G
ti*****ang10	ba****01
y****o13	112********BNQSL

如果恶意监听者选择了更加有利的环境（比如讲座开始之前）进行监听，那么，毫无疑问，他将会获得更多的数据。因此，校园无线网络的安全亟待改善。

① WireShark 支持对截获数据进行过滤，过滤器的书写格式，参见 [4]

第 4 章　解决方案

在本章，我会给出两套解决方案，这两套解决方案是针对 OSI 模型中不同的层次进行的，实现起来的困难程度也不尽相同，但是二者的思想都是一致的：即通过对数据的非对称加密，达到使窃听者无法还原窃听到的数据的目的。当然，这些方案还仅仅是概念性的。由于我的能力有限，掌握信息不够全面，所设想的这些方案在实际实现的过程中是否会遇到意料之外的困难，这是我无法判断的。

4.1　安全意识

页面的编程人员要树立起良好的安全意识，如果没有充足的安全意识，那么一切的加密措施都不能避免漏洞的出现。比如，我们发现的，将记住的账户名和密码保存在 cookie 里，这个问题就不是技术性的问题，这纯属编程人员的疏忽大意所致。因此，最重要的，也是最根本的，是要加强编程人员的安全意识，首先排除这种较为低级的错误。

4.2　HTTPS

这种解决方案是针对应用层的。我们知道，我们能够截获账号和密码是由于 HTTP 协议采用了明文传输。如果能够设法使得 HTTP 数据通过一定的加密之后再进行传输，那么，就会阻止对账号和密码的监听。因此，只需要将登录验证过程的 HTTP 通信转入 HTTPS 协议[①] 即可实现登录过程的保密性的要求。目前我校的信息门户、网络学堂、选课系统等一系列的网站已经采取了这种措施来防止敏感信息泄漏。

① HTTPS 协议实际上是将 HTTP 数据通过 SSL 进行传输，SSL 处于 HTTP 协议之下。因此，整个过程对于 HTTP 协议来说是透明的，这样，登录验证程序基本上无需加以修改。只需要对 HTTP 服务器软件做出调整。

4.3 WPA

但是，基于 HTTPS 协议的解决方案只能治标，不能治本。将校园网登录认证页面转入 HTTPS 协议只能使这一个页面免受监听，却无法对所有用户的所有网络通信数据加密。因此为了保证校园无线网络的更进一步的安全性，学校相关部门应该逐步对设备进行改造，调整设置，使校园无线网络由开放式转为有加密保护的模式，这样一来，可以从根本上解决无线网络窃听的问题，同时也使得一些原本不安全的没有加密的数据免于被窃听和泄密。

插图索引

表格索引

参考文献

[1] Zimmermann H. OSI reference model–The ISO model of architecture for open systems interconnection. Communications, IEEE Transactions on, 1980, 28(4):425–432

[2] http://baike.baidu.com/view/239592.htm

[3] 李海林, 王美琴, 高振明. 基于 Linux 的 802, 11b 无线局域网数据包捕获方法 [J]. 计算机应用研究, 2004, 21(12):270–272

[4] http://www.wireshark.org

[5] Fielding R, Gettys J, Mogul J, et al. Hypertext transfer protocol–HTTP/1.1, 1999

[6] Flanagan D. JavaScript: the definitive guide. O'reilly, 2011

致　谢

感谢×××同学，是他首先发现了校园无线网的安全性问题。感谢×××同学允许我利用他的发现完成这篇论文，我对于没能和他共同完成这篇论文感到十分的遗憾。

感谢×××老师，×老师十分耐心而又细致地指导了我的摘要的英文翻译。

感谢那些与我不曾相识的，被我截获账号名和密码的同学们，如果没有这些一手的实验数据，我的论文就失去了事实上的证据。

感谢 ThuThesis，它的存在让我的论文写作轻松自在了许多，让我的论文格式规整漂亮了许多。

声　明

本人郑重声明：所呈交的论文作业，是本人独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文作业的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

在本文写作过程中因实验需要而截获的隐私信息皆已妥善处理，本人无意也没有利用这些信息进行与学术研究无关的活动。

签　名：_____ 日　期：2013年11月25日