

基于周期特性识别的 WiFi 接入点低耗检测算法

李艾静¹ 董超¹ 陶冰阳² 田畅¹ 王海¹ 皋伟³

(解放军理工大学通信工程学院 南京 210007)¹ (上海交通大学计算机科学与工程系 上海 200120)²
(73841 部队 滨海 210003)³

摘要 持续搜索环境中的 WiFi 信号会严重耗费无线网卡的能量,从而影响靠电池供电的智能手机等移动设备的性能。针对这一问题,利用低能耗的次级设备来识别由 WiFi 接入点广播的周期性信标,以此估计环境中存在的接入点数量与信号强度,并仅当检测到较强网络信号时才唤醒无线网卡进行通信,从而达到节能效果。为了识别信号的周期特性,提出了一种稳定精度折叠算法 PSFA。实验证明,该算法能以较低的耗费识别出稀疏序列中的周期信号,在单接入点和多接入点环境下均可达到较高的识别精度。

关键词 无线网络,热点识别,周期检测

中图分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.3.006

Low-cost Access Point Detection Algorithm Based on Periodicity Identification

LI Ai-jing¹ DONG Chao¹ TAO Bing-yang² TIAN Chang¹ WANG Hai¹ GAO Wei³

(College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China)¹

(Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200120, China)²

(73841 Troops, Binhai 210003, China)³

Abstract Continuous search for access point may consume enormous energy, especially for battery-powered mobile devices like mobile phones. To solve this problem, second device was recommended to search for APs and evaluate their power by detecting the periodic beacons transmitted by APs. Only when WiFi signals are detected can wireless adapter be woken up. To detect periodic signals, PSFA (precision-stable folding algorithm) was proposed. Experiment results show that PSFA can achieve high accuracy under both single AP and multiple APs circumstances with low complexity.

Keywords Wireless networks, AP detection, Periodicity detection

1 引言

随着无线局域网技术的快速发展,智能手机和平板电脑等小型移动设备可以随时随地通过广泛部署的 WiFi 接入点 (Access Point, AP) 接入互联网,以获取所需的服务。而伴随用户对无线上网需求的增加,移动设备快速接入最优网络的能力成为影响用户体验的重要因素。目前无线网卡检测热点的方法是通过搜索接入点发送的信标帧 (Beacon Frame) 来实现的。信标帧是 WiFi 接入点周期性发送的标识自身存在的管理帧,包含了服务集标识 SSID (Service Set Identifier)、时间戳、信标间隔、传输率、负载信息等关于该接入点的基本信息。无线网卡通过不停地搜索并解码信标帧获得周围存在的可用的无线网络信息。但是,由于这些移动设备均靠电池供电,不间断的搜索操作会耗费大量电量,尤其是当周围不存在可用接入点时,持续搜索无线信号会对电池电量造成极大浪费^[1]。

因此,在对周围网络环境未知的情况下,移动设备如何快速、低耗地检测环境中存在的接入点,并估计不同网络的信号强度,具有十分重要的作用。

智能手机之类的移动设备由于已配置了各种传感器,因此可以利用这些低耗传感器来探测网络状态,仅当发现较强网络信号时才开启无线网卡,否则网卡处于休眠状态,以此降低设备能耗。利用传感器设备来检测环境中的接入点情况具有两方面的挑战:1) 传感器的采样频率远低于无线网卡, WiFi 通信需要的带宽为 20M,而多数传感器仅支持 2M 的采样频率,因而无法完整获取信标分组;2) 传感器通信的调制解调和编解码方式与 WiFi 不同,所以无法通过解码操作来识别信标帧并获取进一步的信息。因此,要想利用低耗的传感器来实现 WiFi 接入点的检测,如何从欠采样信号中识别出信标并根据识别结果推断环境中存在的接入点信息是需要解决的主要问题。

到稿日期:2014-04-25 返修日期:2014-07-13 本文受国家自然科学基金项目(61103224,61371124),江苏省自然科学基金项目(BK2011118)资助。

李艾静(1986—),女,博士生,主要研究方向为无线网络、认知网络,E-mail:lishan_wh@126.com;董超(1980—),男,博士,副教授,CCF 会员,主要研究方向为网络编码、Ad hoc 网络、认知网络;陶冰阳(1991—),男,硕士生,主要研究方向为无线网络和博弈论;田畅(1963—),男,博士,教授,主要研究方向为网络编码、无线路由、数据融合;王海(1972—),男,博士,教授,主要研究方向为网络编码、Ad hoc 网络、无线路由;皋伟(1979—),男,工程师,主要研究方向为无线网络、无线通信。

本文利用信标帧的周期性广播特点,提出利用识别周期性信标的个数来推断接入点的个数的新思路,并设计一种基于稀疏序列的稳定精度折叠算法(Precision-Stable Folding Algorithm, PSFA)。PSFA 算法可以快速识别出离散稀疏序列中的周期信号,因而可以通过判断欠采样序列中周期性信号的个数和信号强度来检测环境中存在的接入点个数以及接入点的信号强度,从而避免了对欠采样序列进行解码解调,解决了利用采样率较低的传感器来探测宽带无线网络信号的问题。相较于信标解码,周期信号的识别对采样频率和设备运算能力的要求大大降低,从而降低了移动设备的能量耗费。

2 相关工作

Shih^[2]首先提出利用低能耗次级无线电设备替代无线网卡与接入点通信来发现连接机会、降低能耗。紧随其后,多种类似协议被陆续提出,如 Cell2notify^[3]、Coolspots^[4]和 E-sense^[5]。这些研究的网络发现机制依然基于对信标帧的解码完成。根据次级设备获取网络信息的方式,这些研究主要可以分为两类:一类是次级设备被动接收信标,从信标解码获取网络信息;另一类是次级设备主动与接入点交互,但这需要接入点的配合,要对现有基础设施做出大量修改。针对上述研究的缺陷,本文提出一种利用周期信号识别的方法获取网络信息,这种方法既不需要对信标帧进行解码,也不需要对接基础设施做出任何修改。

离散序列的周期检测方法在信号处理领域得到了广泛的研究,目前采取的主要方法有离散傅里叶变换^[6]和自相关运算^[7],但是这两种方法并不适用于本文场景。首先,信标信号的发送并不遵循严格的周期,接入点有数据正在发送时会等待数据发送结束后才发送信标信号,这会使信标周期发生毫秒级的偏差;其次,这两类算法仅能检测出周期信号的存在,并无法确切计算周期信号的个数和准确的周期。此外,大多无线网络的信标周期设置在 100ms 左右,为了获取足够的采样点以分析其周期性,需要采样 1s 的信号,对于如此多的采样点,离散傅里叶变换和自相关的运算的开销都是巨大的,传感器的运算模块难以承受。本文提出的 PSFA 算法基于稀疏序列运算,而且仅有加法操作,可以极大地降低运算开销。

3 折叠算法

3.1 基本折叠算法

折叠算法^[1,8]是一种用于检验序列中是否存在周期为 P 的信号的算法。算法基本思路是将序列按周期 P 分段,然后将信号段对齐叠加。考虑由 N 个采样点构成的序列 $R[n]$, $R[n]$ 中含有周期为 P 的信号,则折叠后的结果 $F[i]$ 包含 P 个元素:

$$F[i] = \sum_{j=0}^{\lfloor N/P \rfloor - 1} R[i+j \cdot P], i=1, \dots, P \quad (1)$$

在正确的周期处折叠后,周期信号会因叠加而增强,得到一个较高的峰值,而噪声和干扰信号叠加后只会得到幅度很小的零散峰值,因此可以很好地将周期信号从序列中区分。图 1 给出了一个简单的示例,序列在位置 2 处存在周期为 5 的信号,叠加后,在位置 2 处得到峰值 4,而其它位置得到的叠加值很小。

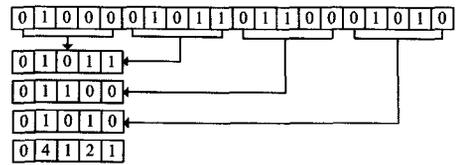


图 1 基本折叠算法

由于信标的周期未知,因此需要对所有可能的周期进行计算(约为 251000ms,以 1ms 为步长变化),所以,基本折叠算法的运算量为:

$$\sum_{P_i \in P} (N - P_i)$$

而实际上,目前的商用设备默认的信标周期均为 100ms,受时钟影响,实际的信标周期为 102.4ms(Wireshark 抓包实际间隔)。基本折叠算法只能处理整数周期,若将小数周期截断为整数,会导致叠加时周期信号无法对齐,进而无法得到较高的叠加峰值。

3.2 PSFA 算法

为了解决基本折叠算法处理小数周期准确性不高的问题,本文提出一种基于稀疏序列的稳定精度折叠算法 PSFA。

3.2.1 预处理

从图 2 中对 1s 时间内 AP 发送数据的采样可以看出,信标信号稀疏地分布于时域采样点中。为了便于计算,首先对采样序列进行预处理。预处理主要包括去噪和量化操作。经过处理的信号,其噪声和数据采样值变为 0,而信标帧采样值变为 1,由此,原始采样序列转化为稀疏 0-1 序列。

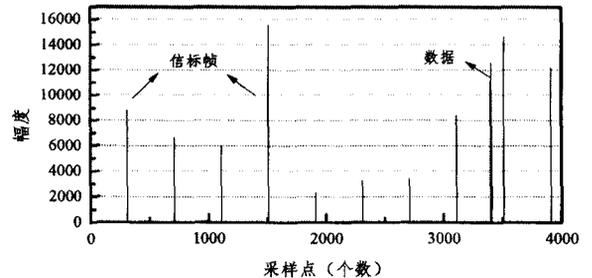


图 2 AP 信号采样示意图

噪声的去除通过预设门限完成,只有高于一定能量门限的采样点才被设置为 1,而能量较低的采样点以及噪声设置为 0。为了保证唤醒后的网卡可以解码数据,门限的设置至少应该保证数据解码所要求的最低信噪比,实验中采用 15dB。除了噪声之外,数据信号也会引起折叠结果的不准确,因此,数据信号也要从采样序列中去除。根据 IEEE 802.11 协议,信标长度可以为 8200 字节,因此其持续时间为 256 μ s 到 1720 μ s 之间,而持续时间在此区间之外的信号可以认为不是信标帧,将这些采样点的值置为 0。采样频率的设置必须满足所有的信标信号至少被采样一个点,因此,采样频率应不低于 4k/s。相较于传统方式,周期检测所需的采样率 (4k/s) 仅为信标解码 (20M) 的 1/5000,大大降低了对检测设备处理能力的要求和能量的耗费,因而可以利用传感器设备进行周期信号的识别。

3.2.2 PSFA 算法流程

PSFA 算法是对基本折叠算法的改进,可以对稀疏序列进行非整数周期的快速检测。

PSFA 算法的计算思路与基本折叠算法类似,但是处理

的是小数周期。图 3 给了出 PSFA 算法的一个实例,信号周期为 3.6,折叠结果中包含 4 个元素。图中存在大量形如“A”、“B”,在折叠结果中跨越两个元素的采样信号,这些采样点对于折叠结果的两个元素均有贡献,因此,需要定义每个采样值对折叠结果每个元素的贡献。

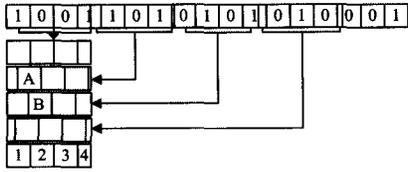


图 3 PSFA 算法

对于采样序列中的每个元素 $R[i]$,定义其偏移量为:

$$\Delta_i = i - \lfloor i/P \rfloor \cdot P \quad (2)$$

设

$$u_i = \lfloor \Delta_i \rfloor \quad (3)$$

$$r_i = \Delta_i - u_i \quad (4)$$

则第 i 个采样点对折叠结果第 k 个元素的贡献值定义为:

$$C_{i,k} = \begin{cases} 1 - r_i, & k = u_i \\ r_i, & k \bmod \lceil P \rceil = (u_i + 1) \bmod \lceil P \rceil \\ 0, & \text{其它} \end{cases} \quad (5)$$

加权贡献可以表示为:

$$Contrib_{i,k} = R[i] \cdot C_{i,k} \quad (6)$$

则 PSFA 折叠后的结果 $F[k]$ 包含 $\lceil P \rceil$ 个元素,其中

$$F[k] = \sum_{i=1}^N Contrib_{i,k} \quad (7)$$

从式(6)可看出,若 $R[i]=0$,则对于任意 k ,有 $Contrib_{i,k} = 0$ 。而采样信号在经过预处理之后,原始采样序列成为存在大量 0 的稀疏序列。利用稀疏序列,可以将算法的加法运算次数由 $N-P$ 降为 γN ,其中 γ 为信号密度,定义为:

$$\gamma = \frac{|\{R[i] | R[i]=1\}|}{|R|} \quad (8)$$

当对多个可能的周期进行搜索时,可以通过指定步长精度来实现。由于步长精度可以根据实际周期的取值而指定(本文中采用精度为 0.1),因此 PSFA 算法对周期检测的精度是稳定的。

3.2.3 迭代检测

由于环境中可能存在多个无线网络信号,因此要检测出接入点的个数,需要对采样信号进行迭代运算。当检测到一组周期信标后,需要将其从采样点中去除,然后对剩余信号继续搜索周期信号,直到所有的周期信号被识别。

迭代检测的流程参见算法 1。迭代运算之后,采样序列中所有的周期信号被识别出来,同时返回这些信号的周期及其平均能量,由此,次级设备可以判断当前环境中存在的热点个数及其信号强度,从而进一步决定是否唤醒网卡。

算法 1 迭代 PSFA 周期检测算法

Input: R_{pre} 即预处理后采样序列

P 即信标周期集合

α 即门限

Output: P_{out} 即检测出的信标周期

E_{out} 即检测到的周期信号的平均能量

Process:

1. eliminate data from R_{pre}
2. while true do
3. Compute the folding result $F_{p[i]}$ of each $P[i]$

4. // find the max normalized folding result of each R_{pre}
5. $F_{max} = P[i] \cdot \max(F_{p[i]})$
6. sort F_{max} in the descending order, as well as corresponding P and F_p
7. if $F_{max}[0] \geq \alpha$
8. $P_{out}[k++] = F_{max}[0]$
9. remove detected beacons from R_{pre}
10. else
11. break
12. end if
13. end while

4 实验与分析

4.1 准确性分析

4.1.1 单 AP 环境下的检测

实验采用 D-Link 无线路由器建立无线连接热点,配有 Atheros AR5001X+无线适配器的台式机作为用户设备,用户与热点之间建立 FTP 连接,通过调整业务流量设置不同的信道利用率,利用通用软件无线电外围设备 USRP 实时采集环境中信号,并对采集到的信号进行在线分析。USRP 是国内外使用最为广泛的软件无线电平台,其实现代码可以方便地移植到传感器等其他设备中。

USRP 以 4k/s 的采样率对环境信号进行采样,图 4 和图 5 分别显示了采样序列在正确的周期和不正确的周期折叠之后的结果。可以看出,按照正确的周期折叠,由于周期信号在同一位置重复出现,因此可以得到一个较高的峰值,而在错误周期处折叠时,周期信号无法对齐相加,因此只能得到零星的幅度较小的值。

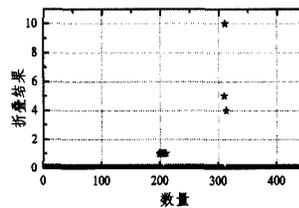


图 4 在正确周期处折叠的结果 (102.4ms)

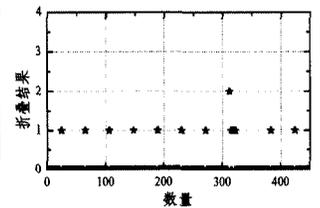


图 5 在错误周期处折叠的结果 (111.6ms)

周期检测的准确性与信道利用率之间的关系如图 6 和图 7 所示。实验比较了不同信道利用率下的虚警概率和漏警概率。可以看出,漏警概率一直维持在较低的水平,即使在信道利用率达到 0.2 时,漏警概率依然不超过 5%,而虚警概率随着信道利用率的增加而增加。通过对 SIGCOMM 2008^[9]大会会场采集的实际无线数据流量进行统计分析可知,实际无线网络的信道利用率平均值约为 7.6%,而此时周期检测算法的虚警概率和漏警概率分别为 0.3% 和 2.3%。

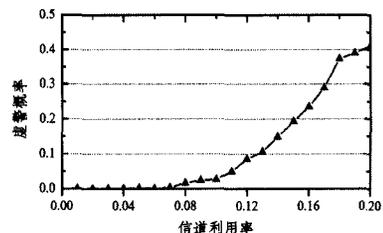


图 6 周期检测虚警概率与信道利用率的关系

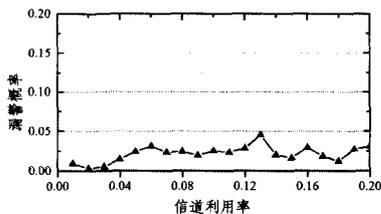


图7 周期检测漏警概率与信道利用率的关系

4.1.2 多 AP 环境下的检测

在实际的无线环境中,一般存在多个热点,因此需要对多个 AP 的信标信号进行识别。本文测试了实验室内存在 3 个 AP 热点时所提算法的性能,实验中,固定 AP1 的信标周期为 100ms, AP2 和 AP3 的信标周期在 [80, 120] ms 区间内变化。总共进行 3 组实验,不同组之间 AP 具有不同的能量差异,第三组为对照实验组。实验参数设置见表 1。

表 1 多 AP 参数设置

	第一组			第二组			第三组
	AP1	AP2	AP3	AP1	AP2	AP3	AP1
能量 (dbm)	-72	-39	-42	-52	-31	-37	-54
信标周期 (ms)	100	[80, 120]	[80, 120]	100	[80, 120]	[80, 120]	100

实验结果如表 2 所列。实验结果显示, AP 数量增多会使检测准确性下降,但是,仍然可以得到 90.4% 以上的准确率。

表 2 多 AP 环境下周期检测准确率

	虚警概率	漏警概率	准确率
第一组	3.57%	6.02%	90.4%
第二组	2.11%	4.08%	93.8%
第三组	0.00%	0.00%	100%

为了进一步验证算法对实际环境中多接入点的检测性能,我们对上海交通大学电信楼实验室环境下的 WiFi 信号进行检测。利用无线网卡和无线网络搜索软件检测出环境中存在 9 个热点,如图 8 所示。其中包括各个实验室无线热点 6 个、校园网热点 1 个,以及公共热 2 个 (CMCC-EDU、ChinaUnicom)。这些热点工作于 3 个完全正交的信道上,工作在同一信道的热点数量最多为 3 个,利用 PSFA 算法对信道 11 进行检测,识别准确率可以达到 90.7%,检测到的信标周期均为 102.4ms,这与我们在实验室内部自己构建的实验环境上的测试结果是一致的。同时,在无线热点更加密集的场所,由于 AP 之间会通过频段来处理信道争用问题,处于某个相同频段的 AP 数目并不会很多,本文所提算法的准确性依然可以得到保证。

Network Name	BSSID	Security	Protocol	RSSI	Noise	Channel	Band	Width
CCFLAB	c0:61:18:0b:76:3c	WPA/WPA2 Personal	802.11b/g/n	-38	-83	1	2.4GHz	40 MHz
ChinaUnicom	6c:f3:7f:36:3f:50	Open	802.11b/g/n	-73	-92	11	2.4GHz	20 MHz
Qoudteam	f0:f7:53:8b:c4:18	WPA/WPA2 Personal	802.11b/g/n	-57	-92	6	2.4GHz	20 MHz
CMCC-EDU	6c:93:7f:36:3f:54	Open	802.11b/g/n	-73	-92	11	2.4GHz	20 MHz
epcc-fip	29:2c:b2:3b:a8:92	WPA/WPA2 Personal	802.11b/g/n	-63	-83	1	2.4GHz	40 MHz
epcc-print	78:a1:06:bf:5f:10	WPA/WPA2 Personal	802.11b/g/n	-44	-92	6	2.4GHz	40 MHz
Lab-328	64:9e:f3:87:17:2c	WPA2 Personal	802.11b/g/n	-43	-92	11	2.4GHz	20 MHz
MI-WIFI	d8:42:ac:fc:3b:16	WPA/WPA2 Personal	802.11n	-53	-83	1	2.4GHz	40 MHz
SITU	6c:f3:7f:36:37:68	WPA2 Enterprise	802.11a/n	-82	-92	153	5 GHz	20 MHz

图 8 实际无线环境信道占用情况

4.2 算法开销分析

PSFA 算法和基本折叠算法的时间复杂度均为 $O(|P|N)$,可见 PSFA 算法并没有增加基本折叠算法的复杂性,但其运算量相较于基本折叠算法大大降低。

对于单个周期信号的检测,PSFA 算法的加法运算次数为 γN ,其中, γ 为信号密度,其定义见式(8)。而对于精度为 δ 的周期集合 P ,其运算量为:

$$\frac{|P|}{\delta} \cdot \gamma \cdot N \quad (9)$$

表 3 对比了基本折叠算法和 PSFA 算法的运算开销。计算时采用 2000 个采样点,初始检测周期 P_0 为 80,PSFA 算法的检测精度 δ 为 0.1,比较不同周期集合元素数量 $|P|$ 以及信号密度 γ 下的运算开销。可以看出,当信号密度较低时,PSFA 算法的运算开销远低于基本折叠算法,而实际中,稀疏序列的信号密度一般在 0.03 以下。

表 3 PSFA 算法与基本折叠算法运算开销对比

运算量	基本折叠算法 $\sum_{P_i \in P} (N - P_i)$	PSFA $\frac{ P }{\delta} \cdot \gamma \cdot N$
$P_0 = 80, P = 80, \gamma = 0.015$	1.5×10^5	2.4×10^4
$P_0 = 80, P = 80, \gamma = 0.1$	1.5×10^5	1.6×10^5
$P_0 = 80, P = 160, \gamma = 0.015$	3×10^5	4.8×10^4

结束语 本文提出利用识别接入点发送的信标的周期性来估计环境中 WiFi 接入点的数量及能量,并提出 PSFA 算法来识别周期信号并确定其周期。PSFA 算法仅对采样序列进行分析,不需要对信标进行解码,对于采样频率和运算能力的要求也大大降低,因此能够应用于传感器设备中,从而减少移动设备中无线网卡产生的能耗。需要指出,本文提出的 PSFA 算法不仅仅适用于 WiFi 接入点的探测,对于其它具有周期性特征的信号的识别同样适用,算法具有较好的扩展性。

参考文献

- [1] Zhou R, Xiong Y, Xing G, et al. Zifi: wireless LAN discovery via ZigBee interference signatures [C] // Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking. ACM, 2010: 49-60
- [2] Shih E, Bahl P, Sinclair M J. Wake on wireless: an event driven energy saving strategy for battery operated devices [C] // Proceedings of the 8th Annual International Conference on Mobile Computing and Networking. ACM, 2002: 160-171
- [3] Agarwal Y, Chandra R, Wolman A, et al. Wireless wakeups revisited: energy management for voip over wi-fi smartphones [C] // Proceedings of the 5th International Conference on Mobile Systems, Applications and Services. ACM, 2007: 179-191
- [4] Pering T, Agarwal Y, Gupta R, et al. Coolspots: reducing the power consumption of wireless mobile devices with multiple radio interfaces [C] // Proceedings of the 4th International Conference on Mobile Systems, Applications and Services. ACM, 2006: 220-232
- [5] Chebrolu K, Dhokne A. Esense: communication through energy sensing [C] // Proceedings of the 15th Annual International Conference on Mobile Computing and Networking. ACM, 2009: 85-96
- [6] 奥本海姆, 谢弗. 离散时间信号处理 (第 3 版) (英文版) [M]. 北京: 电子工业出版社, 2011: 623
- [7] 胡广书. 数字信号处理导论 [M]. 北京: 清华大学出版社, 2006: 26
- [8] Staelin D H. Fast folding algorithm for detection of periodic pulse trains [C] // IEEE Proceedings. 1969, 57: 724-725
- [9] Sigcomm 2008 traces [OL]. http://www.cs.umd.edu/projects/wifidelity/sigcomm08_traces/