ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 30,No. 2,Mar. 2015,pp. 441－451
DOI:10.16337/j.1004-9037.2015.02.024

http://sjcj.nuaa.edu.cn
E-mail:sjcj@nuaa.edu.cn
Tel/Fax:＋86-025-84892742

基于时间序列分解的用户行为分析

常慧君[1]　单　洪[1]　满　毅[2]　毛　毛[3]

(1.电子工程学院网络系,合肥,230037;2.中国人民解放军94362部队,青岛,266000;3.西昌卫星发射中心,西昌,615000)

摘　要:提出一种用户报文时间序列分解的方法。首先对信息时间序列进行采样,利用不同类型采样信号经低通滤波器后衰减程度不同的特性,用低通滤波器发现和提取序列突发成分;然后基于向量之间的欧几里得距离,用遍历和匹配方法提取周期子序列;最后将报文序列分解为突发成分、周期成分和随机成分。该方法无需解析报文内容,分解出的周期成分可以用来分析用户一般性行为,突发成分可以用来检测突发异常。仿真实验结果表明,该方法能够有效区分序列中的各类成分。

关键词:用户行为分析;时间序列;低通滤波;周期序列提取

中图分类号:TP393　　**文献标志码**:A

User Behavior Analysis Based on Decomposition of Time-Stamp Sequence

Chang Huijun[1], Shan Hong[1], Man Yi[2], Mao Mao[3]

(1. Department of Computer Network, Electronic Engineering Institution, Hefei, 230037, China;2. Troops 94362,PLA, Qingdao, 266000, China;3. Xichang Satellite Launch Center, Xichang, 615000, China)

Abstract: A decomposition method of user packet time-stamp sequence is proposed. Firstly, the method samples the time sequence and utilizes a low-pass filter to extract the burst component based on the different attenuation characteristics of different types of sampled signals. Then, it uses a traversal and matching method to extract the periodic sub-sequence based on Euclid distance between vectors. Finally, it decomposes the encrypted packet sequence into the burst, periodic, and random components. The method does not need to parse the payload of packet, while the periodic component can be used to analyze the user′s routine behavior, and the burst component can be used for burst abnormality detection. Simulation results show that the method can effectively distinguish different components in the sequence.

Key words: user behavior analysis; time-stamp sequence; low-pass filter; periodic sequence extraction

引　　言

用户行为分析是指用某些特征量的统计特征或者特征量的关联关系定量或定性的表示网络用户行为的规律。通过掌握用户行为规律,得以控制并预测网络用户行为。用户行为分析还可用于发现异常用户行为,对网络安全管理有着重要意义。

用户行为最终通过承载用户业务的数据流来体现,而数据流的实质是一段时间序列,因此对用户行

收稿日期:2013-12-25;**修订日期**:2014-05-28

为的分析实质上是对时间序列的分析。对时间序列的分析方法主要分为 3 类:基于时间序列分解的方法、基于聚类的方法和基于关联分析的方法。基于时间序列分解的用户行为分析方法主要是用来分析互联网骨干链路上的多用户查询、访问的规律性[1-3],为网络资源分配和管理提供依据。这类研究的已知信息是在某个时间点或时间段内的流量、查询量或数据量等统计量,在分析单个用户的行为特征时不能有效发挥作用。基于聚类的方法如文献[4-6]等,根据用户的访问偏好等特性对用户进行分类,主要用于营销领域,也可能被利用来实现针对用户的认知欺骗攻击。基于关联分析的方法如文献[7-9]等,分析网络用户的访问偏好与位置和时间等因素的关系,可以用于对单个用户行为的预测。以上方法都假设网络用户的身份、地址和所发的数据内容等信息已知。当用户对自身的数据内容进行加密时,以上用户行为分析方法将不再适用。为保证通信内容的完整性和机密性,目前很多用户业务都采用了端到端加密机制[10-12],使得在通信业务源和目的之间路径上的任何一个第三方无法解析用户的数据内容。

本文提出一种时间序列分解方法。首先对时间序列进行采样,转换为信号序列,采用低通滤波器发现并提取信号序列中的突发成分;然后用遍历和匹配方法提取剩余时间序列中的周期,从而最终将用户相关的时间序列分解为周期成分、随机成分和突发成分。该方法可以在无需对用户数据进行解密的情况下实现对用户行为的分析,且经实际数据验证,能够取得较好的分辨效果。

1 问题模型

假设在某段时间内检测到一段数据报文,由于端到端加密机制的采用,只能解析得到数据包的 IP 报头,而无法获知具体的数据内容。根据 IP 报头内容,提取以某个节点为源地址的所有数据包,将收到数据报文的时刻定义为该报文的时间戳,则该节点发出的数据报文可以表示为一个时间序列 $A=(a_1, a_2, \cdots, a_n)$。

与用户行为相关的时间序列可以分解为突发成分、周期成分和随机成分。其中突发成分可以用来发现用户的异常行为或突发状况,主要体现为在短时间内发送大量数据包;周期性成分主要用于位置和态势信息的更新,体现了用户行为在时间域上的规律性,主要表现为数据段开始时间的周期性;随机成分则是没有明显特征的随机数据流。本文设计的时间序列分解方法将从时间序列 A 中分解出突发子序列、周期子序列和随机子序列,为用户行为的规律性和异常发现提供依据。其中,突发子序列表示为 $\{b_i, e_i, r\}$,即从 b_i 开始到 e_i 结束的一段时间序列,其数据速率 r 大于一定阈值;周期子序列表示为 $\{t_i, T_i, P_i\}$,即以 T_i 为周期,以 t_i 为起始点的周期序列 P_i 的集合,其中 P_i 为一段数据速率大于一定阈值的时间序列,且该阈值小于突发数据的速率阈值;随机成分则为序列中的其余部分。

为便于数学分析,需要对时间序列进行采样,将其转换为时间域上的信号序列。令 Δ 为采样间隔,对 A 中每一个 a_i

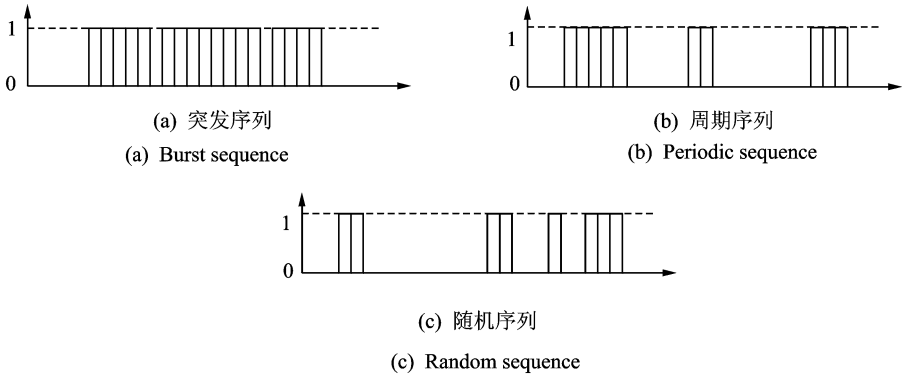
$$f(k\Delta)=\begin{cases}1 & \exists a_i, \text{使得 } k\Delta \leqslant a_i \leqslant (k+1)\Delta \\ 0 & \text{其他}\end{cases} \tag{1}$$

式中:Δ 的值设置为时间序列中最小时间间隔的一半,这样就能保证每个报文的发送时刻都能对应采样序列中的 1。

经过以上采样过程,就将表示数据包时间戳信息的时间序列转换成了一个离散时间函数。在物理意义上,$f(k\Delta)$ 是 (a_n-a_1) 时间内的一组离散时间方波信号,包括突发、周期和随机 3 类子成分。突发成分的采样信号如图 1(a)所示,由于突发数据时间比较集中,相应的采样信号也较为密集,表现为类似矩形脉冲的信号序列。周期成分的采样信号如图 1(b)所示,各数据段的起始时刻服从周期分布,随机成分如图 1(c)所示。

2 基于 FIR 的时间序列突发成分提取

随机发送的单个数据包被采样为单个冲激信号,突发的大量数据被采样为类似矩形的信号,少量的



(a) 突发序列
(a) Burst sequence
(b) 周期序列
(b) Periodic sequence
(c) 随机序列
(c) Random sequence

图 1 不同类型的采样信号示意图

Fig. 1 Illustration of various types of sampled signals

集中数据发送则被采样为较窄的矩形信号。这 3 类信号经过数字低通滤波器后会有不同程度的衰减。其中单个脉冲信号的频谱在整个频域范围内是一个常数,经过低通滤波器后丧失了整个高频频域,因而衰减最大。对于矩形脉冲,若幅度为 E,时间宽度为 τ,其时域表达式为

$$f(t)=\begin{cases}E & |t|\leqslant \tau/2\\ 0 & \text{其他}\end{cases} \tag{2}$$

则其对应的频域表达式为

$$F(w)=E\tau \mathrm{Sinc}(w\tau/2) \tag{3}$$

式中:Sinc(•) 是辛格函数。对应带宽为 $B=1/\tau$,与时宽成反比。因而矩形脉冲的时间跨度越大,经低通滤波器后衰减越小。

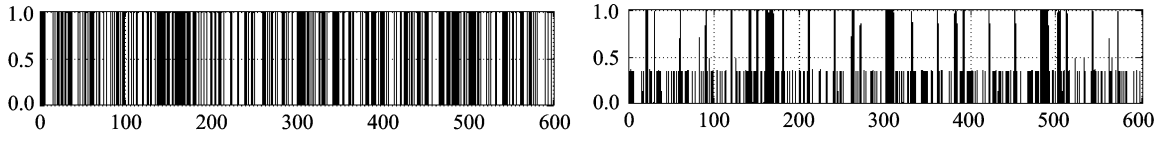
如图 2(a)所示,仿真生成的原始时间序列由突发成分、周期成分和随机成分组成。其中,突发成分分别始于 150,300 和 480 处,周期成分的周期分别为 30 和 120。其采样信号和经 FIR 低通滤波器处理后的信号如图 2(b)所示。可见,宽的突发信号得以较为完整的保存,随机成分则衰减很大。

根据采样信号的以上特性,利用 FIR 滤波器来发现信号序列中的突发成分。FIR 滤波器又称卷积滤波器,是一种数字低通滤波器,其频率响应表达为

$$H(\mathrm{e}^{\mathrm{j}\omega})=\sum_{n=0}^{N-1}b(n)\mathrm{e}^{-\mathrm{j}n\omega} \tag{4}$$

式中:$b(n)$ 为长度为 N 的脉冲响应的抽头系数。

给定采样信号序列,其成分未知,经 FIR 滤波后得到的滤波信号序列,则提取突发信号的原理如下:如果滤波信号序列中第 i 到 j 个采样间隔内的子序列存在 m 个采样值都大于特定门限值,且该段序列的数据发送速率大于某特定阈值,则认为 i 到 j 之间的子序列为突发序列。其中,数据发送速率计算方法为 $m/(j-i+1)$。



(a) 原始时间序列采集信号
(a) Original time sequence sampling signals
(b) FIR滤波信号
(b) FIR filtered signals

图 2 仿真信号序列经 FIR 低通滤波的结果

Fig. 2 Low-pass filtered results of simulated signal sequences

综上所述,对时间序列中突发成分的提取流程如图 3 所示。对采样信号进行基于 FIR 的低通滤波,然后对滤波信号进行突发信号提取,最后对突发信号序列进行逆采样过程,即得到时间序列中的突发成分。

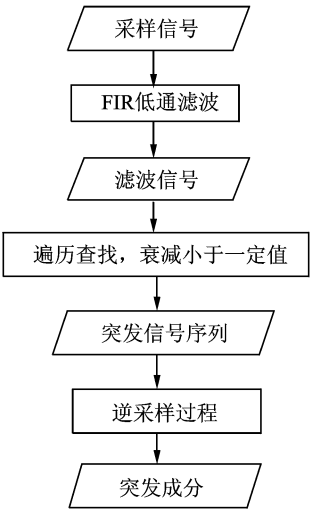


图 3 基于 FIR 的突发成分提取

Fig. 3 Burst component extraction based on FIR

3 基于欧几里得距离的序列周期成分提取

提取出序列中的突发成分之后,得到的信号序列 f 包括周期成分和随机成分。在实际应用中,用户周期性数据每次发送可能不止一个数据包,而是由多个数据包组成的数据段,在提取周期数据发送时间时应当能够提取整段数据的发送时间。这种周期序列相应的采样信号表现为周期性的幅值为 1 的段,实验表明,使用傅立叶变换方法或者功率谱方法不能有效获取这种信号序列的周期信息。因而考虑将采样信号序列 f 重新转换为时间序列 f',直接从该时间序列中提取周期子序列。f 转换为 f' 的方法为:对 f 中每一个值为 1 的采样点,找到 A 中与之距离最近的时刻,最终形成相应的时间序列 f'。用同样的方法可以将 R 转换为突发时间序列 $\{b_i, e_i, r\}$。

本文提出一种基于欧几里得距离的周期序列提取方法,如图 4 所示。其基本原理为:对每一个可能的周期值 T_i 和起始时间 $t_i(1 \leqslant t_i \leqslant T_i)$,生成一个周期序列 B(理想周期序列);对 B 中每一个值 $B(i)$,找到序列 f' 中与之最接近的 $C(i)$,如果 B 和 C 之间的欧几里得距离小于一定阈值,则认为序列中存在以 t_i 为起始时间和以 T_i 为周期的子成分。再以数据发送速率为约束条件,对 C 中的每一个元素,提取可能的一段信息发送时间。其中,由于一般数据的发送速率小于突发数据,算法采用另一个阈值来限制对数据率的取值。

C 和 B 的欧几里得距离计算方法如下

$$D = \sum_{p=0}^{\lfloor a_n/T_i \rfloor} [B(p) - C(p)]^2 \tag{5}$$

其中 D 代表了测试序列 C 与理想周期序列 B 之间的距离。由于信息处理时延的存在,即使是周期发送的时间序列,其与理想周期序列之间的距离 D 必定大于 0。综上所述,时间序列分解模型的总过程如图 5 所示。首先对时间序列进行采样;经过 FIR 滤波器计算,从采样信号中得到衰减程度不同的滤波信号。根据一定的规则约束,可以从滤波信号中提取出突发信号,继而得到对应的突发时间序列。从原始

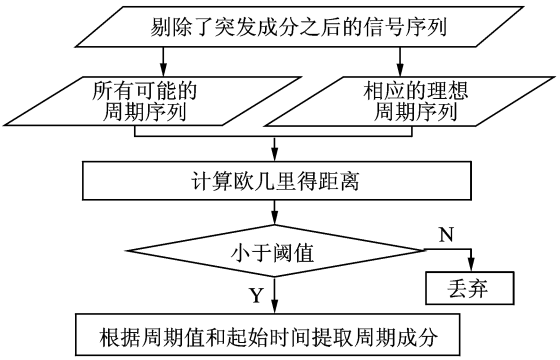


图 4 基于欧几里得距离的周期成分提取

Fig. 4 Periodic component extraction based on Euclid distance

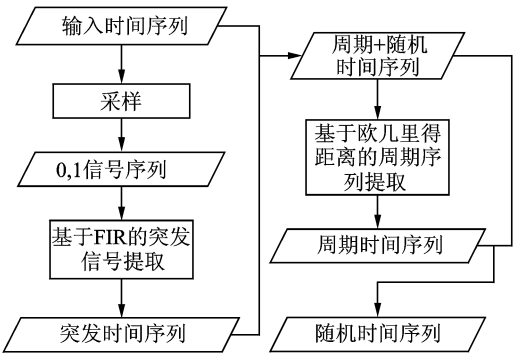


图 5 时间序列分解模型

Fig. 5 Time sequence decomposition model

时间序列中剔除突发序列,采用基于欧几里得距离的方法提取其中的周期序列,最后留下随机序列。

4 仿真分析

仿真实验分为 3 组,第 1 组实验检验算法对序列成分的分辨能力;第 2 组和第 3 组实验检验算法对序列成分提取的准确性。

4.1 算法分辨能力的检验

麻省理工学院林肯实验室 98 年和 99 年搭建了一个局域网,对该局域网进行了时间跨度为两周的数据收集工作,生成的数据集主要用于入侵检测,同样也可以用于用户行为分析。网络结构如图 6 所示,数据收集工作由 192.168.1.90 和 172.16.112.10 两个节点完成。数据集包括网络中每个报文的源、目的节点、报文长度、接收该报文的时间和协议等信息。另外,由于该数据集仅提供了节点的行为方式,而并未对每个节点的突发、周期信息发送的时刻做详细介绍,在本组实验中,本文无法对时间序列分解进行定量分析,而仅对序列分解所推断出的用户行为与实际用户行为进行定性比较。

利用该数据集进行仿真实验。抽取一段以某个节点为源节点的数据报文在嗅探器上的接收时间序

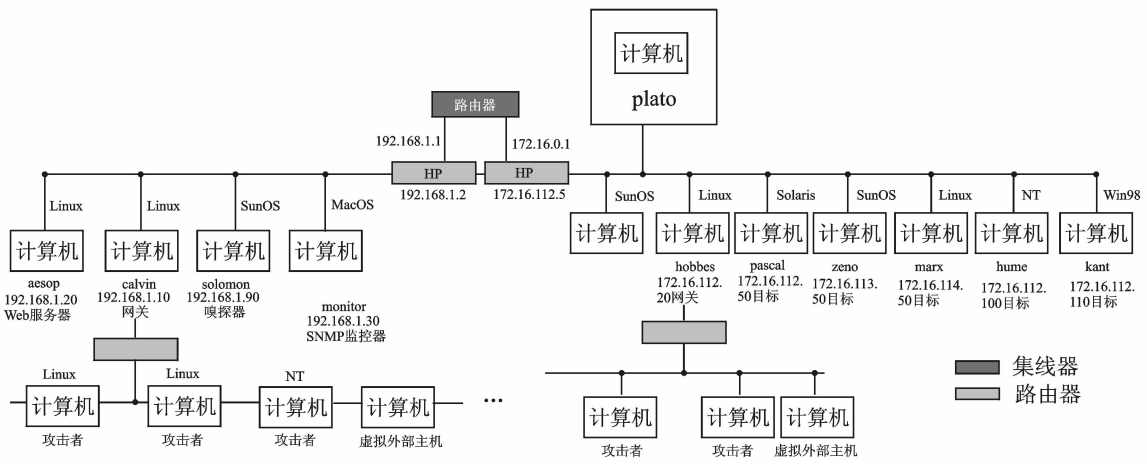


图 6 网络结构示意图

Fig. 6 Illustration of network structure

列，并运行在 Matlab7.0 中实现突发成分提取和周期成分提取，得到该时间序列的突发、周期和随机成分。其中，对时间序列分段处理的主要原因如下：数据集的时间跨度很大（两周），一次性处理会占用大量计算和存储资源，甚至引起资源耗竭；算法的时间复杂度与序列长度的三次方成正比，随着序列长度的增长，算法效率急剧下降。算法主要参数设置如表 1 所示。

表 1 仿真实验参数设置

Table 1 Parameter setting of simulation experiments

参数	数值	参数	数值
FIR 滤波器截止频率	10 Hz	FIR 滤波器阶数	10
数据采样频率	100 Hz	突发滤波信号阈值	0.5
成段数据发送速率阈值	8 Packet/s	欧几里得距离阈值	1.0

以 4 个节点相关的时间序列为例运行分解算法：135.8.60.182，172.16.112.20，196.37.75.158 和 172.16.114.50。其中，135.8.60.182 和 196.37.75.158 为攻击节点，另外两者为正常节点。为保证图片质量和清晰度，仅显示采样信号的前 200 000 个采样点，即时间序列的前 2 000 s。

对 43～19 014 s 之间节点 135.8.60.182 的报文时间戳序列进行采样、低通滤波、周期序列提取的结果如图 7 所示。图 7 中，信号序列的 80%以上由突发子序列组成，因而在节点身份未知的情况下，可以推测该节点存在突发异常行为。

对 37～75 094 s 之间节点 172.16.112.20 的报文时间戳序列运行分解算法的结果如图 8 所示，输入采样信号主要由随机信号组成，说明该节点的信息发送没有明显规律。

对 30 008～39 075 s 之间节点 196.37.75.158 的报文时间戳序列运行分解算法的结果如图 9 所示。图 9 中，输入采样信号由突发信号和一组周期为 272 的周期序列组成，开始时间为 30 046.052 s，相应的时间序列长度为 5 240。

经分析可知，196.37.75.158 与 135.8.60.182 的序列组成不同的原因是其攻击方式的不同。在林肯实验室网站对攻击方式的介绍中，可以看到，196.37.75.158 主要进行了 IP sweep 攻击，是一种侦查扫描攻击方法，该方法需要攻击者不断地采集目标的 IP、端口号等信息，196.37.75.158 采用了一种周期性的采集方式。135.8.60.182 则进行 eject 攻击，是一种缓冲区溢出攻击方法，这种攻击方法需要短期内发送大量特定设置的数据包，远程改变目标节点上的用户权限。

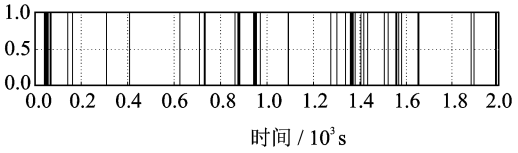
对 14 927～17 862 s 之间节点 172.16.114.50 的报文时间戳序列运行分解算法的实验结果如图 10 所示，输入采样信号主要由突发信号和周期为 159，开始于 14 927.925 2 s，相应时间序列长度为 17 064 的周期序列组成。

节点 192.37.75.158 与 172.16.114.50 的时间序列组成十分相似，但实际上 196.37.75.158 为攻击节点，172.16.114.50 为正常节点。这种情况下，需要利用其他可知因素来对用户行为进行判定，可用信息包括报文类型、报文统计长度或统计包间间隔等。综上所述，本文提出的算法能够区分时间序列中的突发、周期和随机成分，并以此为依据，对用户行为做出正确判断。

4.2 算法准确性的检验

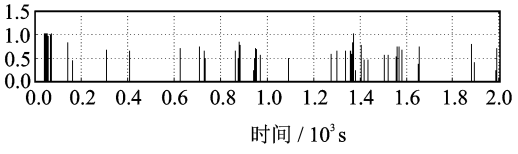
由于林肯实验室数据集的局限性，仅能对算法进行定性分析，在本小节中，通过另一个数据集来检验算法的准确性。

CRAWDAD 网站上提供各类无线网络的追踪数据，用于 MAC 协议，路由协议，业务服务性能或用户行为的研究。其中的 SIGCOMM2009 数据集，是由参加 2009 年在巴塞罗那举行的 SIGCOMM 会议的 76 名人员，利用智能手机上的一种社交应用程序（MobiClique）生成。业务数据包括蓝牙设备接近检测和用户信息交换数据，并在每个设备上记录这些数据的源、目的地址和报文时间戳信息。其中，智能



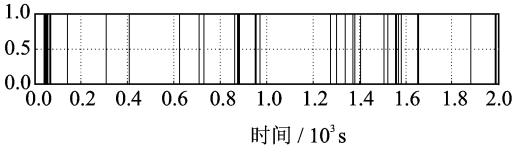
(a) 输入信号

(a) Input signals



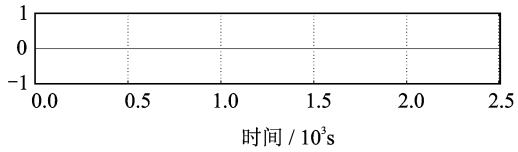
(b) 滤波信号

(b) Filtered signals



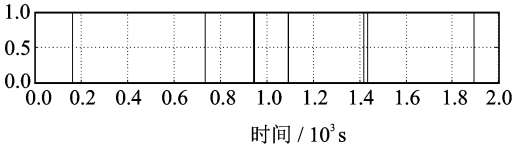
(c) 突发信号

(c) Burst signal



(d) 周期信号

(d) Periodic signal

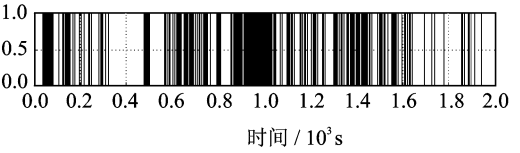


(e) 随机信号

(e) Random signal

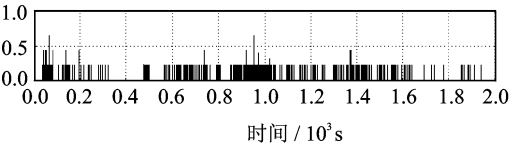
图 7 135.8.60.182 的时间序列采样信号

Fig. 7 Time sequence sampling signals of 135.8.60.182



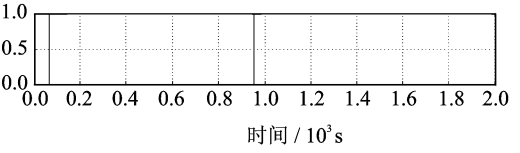
(a) 输入信号

(a) Input signals



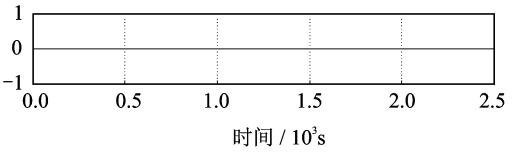
(b) 滤波信号

(b) Filtered signals



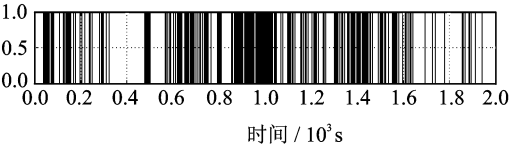
(c) 突发信号

(c) Burst signal



(d) 周期信号

(d) Periodic signal



(e) 随机信号

(e) Random signal

图 8 172.16.112.20 的时间序列采样信号

Fig. 8 Time sequence sampling signals of 172.16.112.20

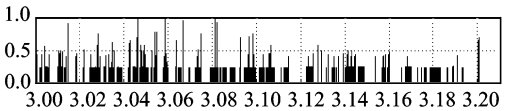
手机的配置为 HTCs620 Windows Mobile 系统,用蓝牙方式彼此传输数据,设备之间的距离一般在 10～20 m。

每个智能手机以(120±10.24) s 为周期发送持续时长为 10.24 s 的蓝牙设备探测信息,搜索周围的蓝牙设备,每发现一个新的可用连接,手机之间都会建立一个 RFCOMM 链路来传输数据。设备之间用 RFCOMM 链路传输会议相关的信息,并在本地记录所发送和接收的信息,记录内容包括每个信息报文

时间 / 10^4 s

(a) 输入信号

(a) Input signals



时间 / 10^4 s

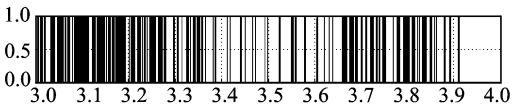
(b) 滤波信号

(b) Filtered signals

时间 / 10^4 s

(c) 突发信号

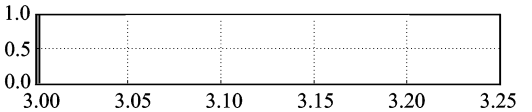
(c) Burst signal



时间 / 10^4 s

(d) 周期信号

(d) Periodic signal



时间 / 10^4 s

(e) 随机信号

(e) Random signal

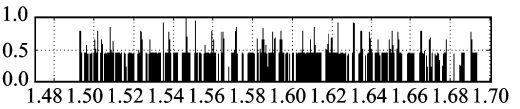
图 9 196.37.75.158 的时间序列采样信号

Fig. 9 Time sequence sampling signals of 196.37.75.158

时间 / 10^4 s

(a) 输入信号

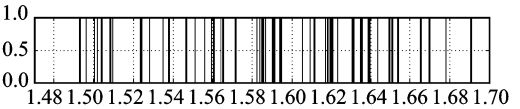
(a) Input signals



时间 / 10^4 s

(b) 滤波信号

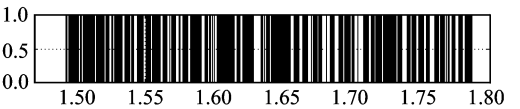
(b) Filtered signals



时间 / 10^4 s

(c) 突发信号

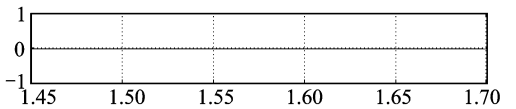
(c) Burst signal



时间 / 10^4 s

(d) 周期信号

(d) Periodic signal



时间 / 10^4 s

(e) 随机信号

(e) Random signal

图 10 172.16.114.50 的时间序列采样信号

Fig. 10 Time sequence sampling signals of 172.16.114.50

的序号、时间戳、源节点 ID、目的节点 ID 和信息类型等。

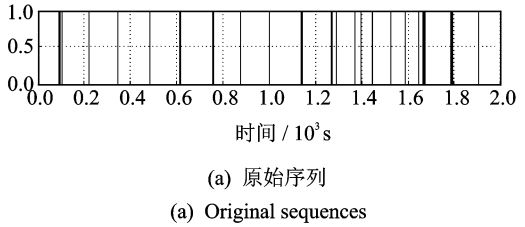
在通信过程中，每个设备上都必定会发送周期性的蓝牙设备探测数据和随机性或突发性的会议相关数据。选取网络中的几个节点，以其信息发送时间序列为输入，运行在 Matlab 上实现的时间序列分解算法，实验参数如表 1。将输出结果与网站中对节点行为的介绍相对比，即可实现对算法的检验。

主要用 3，4 和 9 等节点上的信息发送时间序列来完成检验过程。对节点 3～前 2 000 s 时间内的信息发送时间序列分解的结果如图 9 所示。节点 4 上 30 689～20 3311 s 的信息发送时间序列分解情况如

图 10 所示。节点 9 上 39 531～26 6504 s 之间的信息发送时间序列分解如图 11 所示。

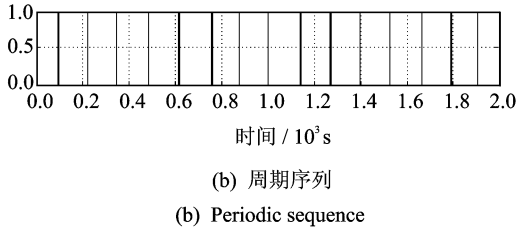
由图 11 可见,节点 3 的信息发送主要由周期成分和随机成分组成,算法运行结果给出的周期值为 130 s,与上文中对节点发送蓝牙设备搜索数据的周期相吻合。

由图 12 可见,节点 4 的信息发送也是由周期成分和随机成分组成,算法运行结果给出的周期序列周期值为 133 s,与(120±10.24) s 误差 3 s。误差是由于用户设备时间不是严格同步,造成每个周期的



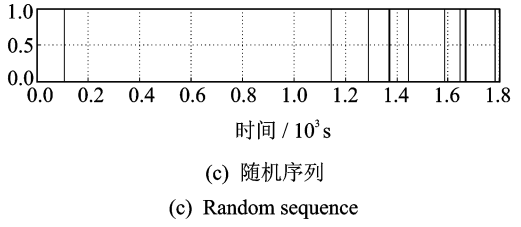
(a) 原始序列

(a) Original sequences



(b) 周期序列

(b) Periodic sequence

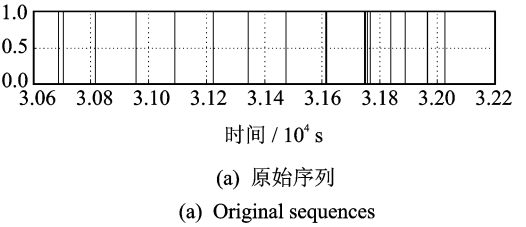


(c) 随机序列

(c) Random sequence

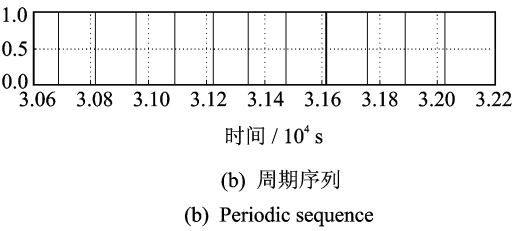
图 11 节点 3 序列成分提取

Fig. 1 Sequence component extraction of node 3



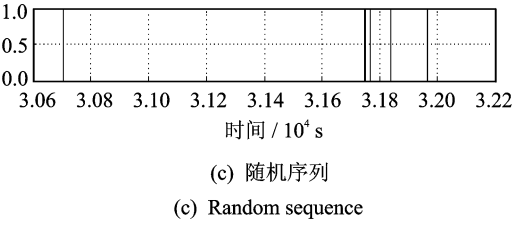
(a) 原始序列

(a) Original sequences



(b) 周期序列

(b) Periodic sequence

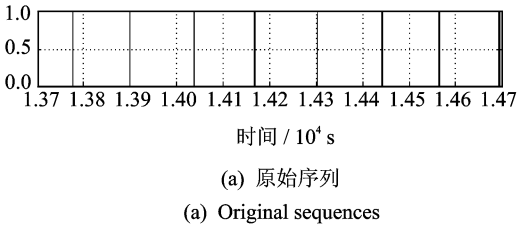


(c) 随机序列

(c) Random sequence

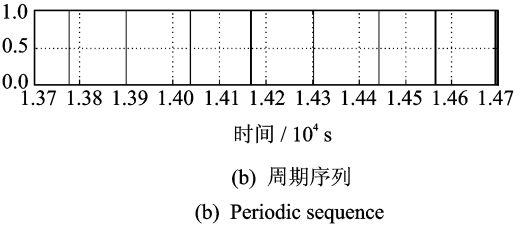
图 12 节点 4 序列成分提取

Fig. 12 Sequence component extraction of node 4



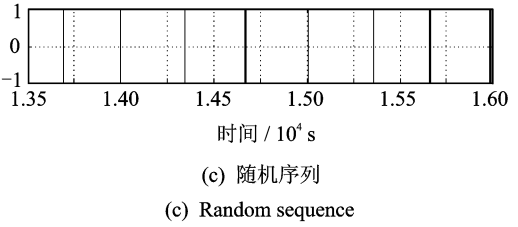
(a) 原始序列

(a) Original sequences



(b) 周期序列

(b) Periodic sequence



(c) 随机序列

(c) Random sequence

图 13 节点 9 序列成分提取

Fig. 13 Sequence component extraction of node 5

用户信息并不一定严格按照(120±10.24) s的要求开始发送。

由图13可见,节点9的信息发送同样仅由周期成分和随机成分组成,算法运行结果给出的周期序列周期值为133 s,与(120±10.24) s的误差在3 s内。综上所述,算法能够准确提取序列中的周期成分。

4.3 通过仿真序列检验算法

公开获取的数据集,由于其应用环境所限,不能完全满足本文中的实验需求。例如,林肯实验室网站提供的数据集中没有提供用户周期信息发送的具体周期值或突发信息发送的开始时刻,SIGCOMM2009数据集中则缺少突发的用户信息发送。因此,在本小节中,仿真生成时间序列来检验本文算法的准确性。

在Matlab平台上模拟生成理想的混合序列,该序列是一段分布在0～600 s区间范围内的时间戳序列,由周期成分、突发成分和随机成分组合而成。周期成分包括两组周期数据,第1组周期数据起始点为0,周期为30 s,即每隔30 s生成一段连续的时间戳,每段时间戳覆盖时长1 s,数据发送率即时间戳之间的时间间隔为0.2;第2组周期数据起始时间为第20 s,周期为120 s,每段覆盖时长3 s,时间戳之间间隔0.2 s。突发成分由3段组成,起始时间分别为160,300和480 s,覆盖时长为9 s,数据发送率为0.02。然后在0～600 s的时间范围内随机生成一组时间戳,随机确定这些时间戳的未知和长度。

以上混合序列作为输入,运行在Matlab平台上实现的时间序列分解算法,包括采样、突发成分提取和周期成分提取方法,仿真参数的设置如表1所示。算法产生的原序列采样、FIR滤波、突发序列、周期序列和随机序列的采样如图14所示。

由图14可见,算法提取的突发成分起始点分别为160,300和480 s,周期成分包括两组,且其周期值与上文介绍相吻合,故本文的时间序列分解算法准确提取了混合序列中的突发成分和周期成分。

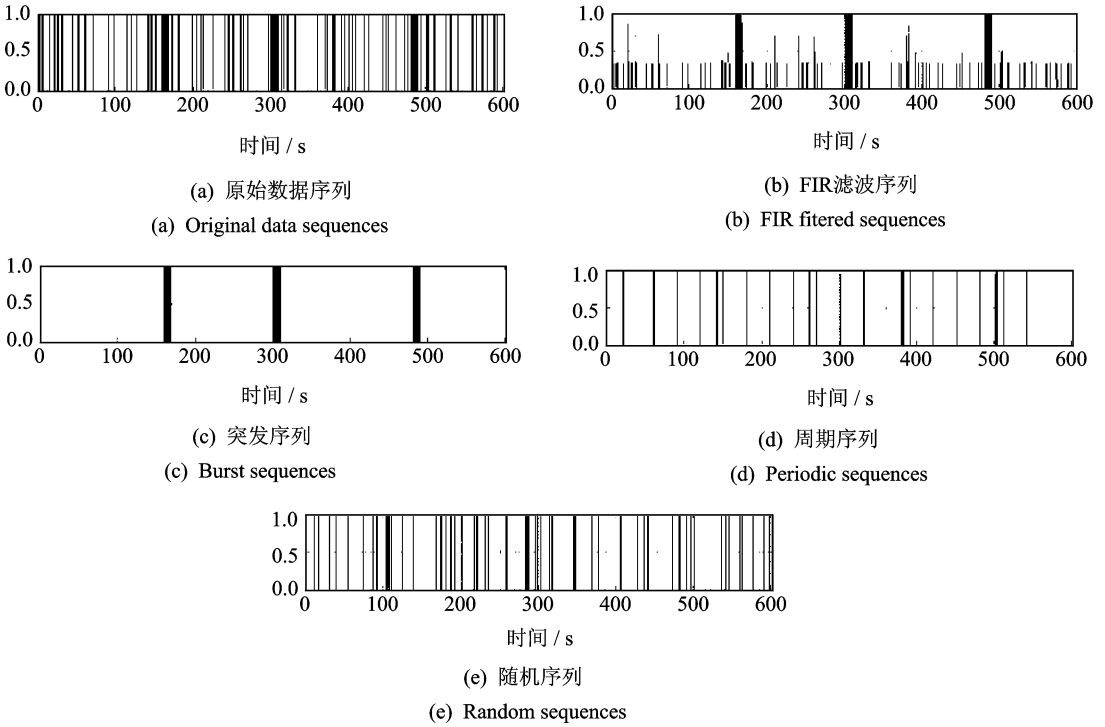


图14 仿真时间序列的成分提取

Fig. 14 Component extraction of simulated time sequences

5 结束语

本文提出了一种用户报文时间序列的分解模型,通过此分解过程,可以获知报文发送行为的突发性、周期性和随机性规律。对公开获取的数据集和仿真生成的时间序列进行仿真实验的结果表明,该方法能够提取出节点发送的数据报文时间序列的突发、周期和随机成分,用于分析用户行为特征。

参考文献:

[1] Vlachos M, Meek C, Vagena Z. Identifying similarities, periodicities and bursts for online search queries[C]// Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data. Paris: ACM, 2004: 131-142.

[2] Radinsky K, Svore K, Dumais S, et al. Modeling and predicting behavioral dynamics on the web[C]// Proceedings of the 21st International Conference on World Wide Web. Lyon: ACM , 2012: 599-608.

[3] Shafer I, Ren K, Boddeti V N, et al. Rainmon: An integrated approach to mining bursty time series monitoring data[C]// Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Beijing: ACM, 2012: 1158-1166.

[4] Wan Li, Liao Jianxin, Zhu Xiaomin. A frequent pattern based framework for event detection in sensor network stream data [C]//Proceedings of the Third International Workshop on Knowledge Discovery from Sensor Data. Paris: ACM, 2009: 87-96.

[5] Qin Tao, Guan Xiaohong, Long Yi, et al. Users′ behavior character analysis and classification approaches in enterprise networks[C]// Proceedings of the 2009 8th IEEE/ACIS International Conference on Computer and Information Science. Shanghai: ACM , 2009: 323-328.

[6] Song Bin, Wang Pingli, Wang Ling, et al. The research on the application of ant colony algorithm at behavior clustering of network users[C]// 2010 International Conference on Internet Technology and Applications. Wuhan: IEEE , 2010: 1-4.

[7] Wu Liang, Chin Aivin, Zhou Yuanchun, et al. Context-aware prediction of user′s first click[C] // Proceedings of the 1st International Workshop on Context Discovery and Data Mining. Beijing: ACM, 2012.

[8] Spiegel S, Gaebler J, Lommatzsch A, et al. Pattern recognition and classification for multivariate time series[C]// Proceedings of the 15th International Workshop on Knowledge Discovery from Sensor Data. San Diego: ACM, 2011: 34-42.

[9] Sridevi S, Rajaram S, Swadhikar C. An intelligent system for time series data using periodic pattern mining in temporal databases[C]// IITM ′10: Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia. Uttar Pradesh: ACM, 2010: 163-171.

[10] Castiglione A, Cattaneo G, De M G, et al. SECR3T: Secure end-to-end communication over 3G telecommunication networks [C]// 2011 15th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS). Seoul: IEEE, 2011: 520-526.

[11] Cheneau T, Sambra A V, Laurent M. A trustful authentication and key exchange scheme for ad hoc networks[C]//2011 5th International Conference on Network and System Security (NSS). Milan: IEEE, 2011: 249-253.

[12] Zhang Wujun, Zhang Yueyu, Chen Jie, et al. End-to-end security scheme for machine type communication based on generic authentication architecture[C]//2012 4th International Conference on Intelligent Networking and Collaborative Systems (INCoS). Bucharest: IEEE, 2012: 353-359.

作者简介:**常慧君**(1986-),女,博士研究生,研究方向:无线网络对抗技术, E-mail :changhj2417@126.com;**单洪**(1965-),男,教授,博士生导师,研究方向:无线网络对抗技术;**满毅**(1985-),女,博士研究生,研究方向:无线网络对抗技术;**毛毛**(1986-),男,硕士研究生,研究方向:无线网络对抗技术。