软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2011,22(Suppl.(2)):17−26

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

一种基于包序重排的流水印技术*

张连成+, 王振兴, 徐 静

(解放军信息工程大学 信息工程学院 网络工程系,河南 郑州 450002)

Flow Watermarking Scheme Based on Packet Reordering

ZHANG Lian-Cheng+, WANG Zhen-Xing, XU Jing

(Department of Network Engineering, College of Information Engineering, PLA Information Engineering University, Zhengzhou 450002, China)

+ Corresponding author: E-mail: liancheng17@gmail.com

Zhang LC, Wang ZX, Xu J. Flow watermarking scheme based on packet reordering. *Journal of Software*, 2011,22(Suppl.(2)):17−26. http://www.jos.org.cn/1000-9825/11023.htm

Abstract: Watermark carriers of existing network flow watermarking schemes are limited to packet payload, traffic rate, and packet timing. However, packet payload is based on flow watermarking schemes, which depend on specific application protocols, such as telnet and rlogin, but encryted traffic and are invisible to traffic interceptors. At the same time, traffic rate and packet timing based ones are vulnerable to timing perturbation introduced by network transmission and attackers. Even worse, most of them have a low watermark capacity and are visible to multi-flow attack, mean-square autocorrelation attack and timing analysis attacks. This paper utilizes packet order as a watermark carrier and proposes a novel packet reordering based flow watermarking (PROFW) scheme. To achieve robustness against packet out-of-order pertubation, a theory of error correcting code is introduced into watermark encoding. Meanwhile, this paper utilizes a stochastic modulation approach to increase the stealthiness of PROFW scheme by controlling packet reordering degree not exceeding normal levels. Empirical results prove its robustness against timing and packet out-of-order pertubations, introduced by network transmission and deliberately by attackers. Compared with typical flow watermarking schemes, PROFW scheme, which has a higher watermark capacity, is more robust against timing and packet out-of-order pertubations.

Key words: network flow watermarking; packet reordering; robustness; error correcting code; stochastic modulation

摘 要: 当前流水印载体局限于包载荷、流速率和包时间 3 种.然而,基于包载荷的流水印技术与具体应用层协议有关,难以处理加密流量,且易被检测和过滤;基于流速率和基于包时间的流水印技术难以从根本上抵御时间扰乱,且存在易被检测、水印容量小等问题.采用包序作为流水印载体,提出一种基于包序重排的新型流水印技术 PROFW.将纠错码理论引入到水印信息编码中,大大提高了 PROFW 技术的鲁棒性,并引入概率调制思想,将包序重排程度控制在正常范围内,保证了 PROFW 技术的隐蔽性.测试结果表明,PROFW 技术在保证隐蔽性的前提下,对于自然产生

* 基金项目: 国家高技术研究发展计划(863)(2006AA01Z449, 2007AA01Z2A1); 国家重点基础研究发展计划(973)(2007CB307102)

收稿时间: 2011-02-15; 定稿时间: 2011-05-31

和主动引入的时间干扰和包乱序具有较强的鲁棒性.与当前典型流水印技术相比,PROFW 技术不但在应对时间扰乱和包乱序时的鲁棒性更强,而且提高了水印容量.

关键词: 流水印;包序重排;鲁棒性;纠错码;概率调制

作为一种主动流量分析(active traffic analysis)手段,流水印(network flow watermarking)技术近年来得到广泛关注,它们通过主动延迟所选数据包或细微改变流量速率等方式,将水印信息嵌入至发送端发送的网络数据流中,然后再在接收端附近对嵌入水印信息的标记数据流中的水印信息进行解码恢复,通过对比恢复出的水印和原始水印之间的相似度来实现数据流关联,进而达到数据流追踪的目的,可广泛应用于匿名用户关联[1−3]、匿名网络电话追踪[4]、跳板攻击源定位[5−9]和僵尸主控机发现[10]等方面.

与仅使用包时间(但不改变)、包数量或包大小等流量特征的被动流关联(flow correlation)方法[11−14]相比,流水印技术准确率更高、误报率更低,且只需要更少的数据包便能更有效地追踪恶意数据流.

然而,已有流水印技术所采用的流水印载体(watermark carrier)局限于包载荷(packet payload)[15]、流速率(traffic rate)[1,2]和包时间(packet timing)[3,8,16,17].为抵御网络传输引入的自然时延和攻击者主动添加的时间扰乱(timing perturbation)等干扰,进而达到较高的检测率,现有流水印技术必须向数据流中引入大的时延,导致攻击者可发起多流攻击(multi-flow attack)[18]、均方自相关(mean-square autocorrelation,简称 MSAC)攻击[19]和时间分析攻击(timing analysis attack)[20]等来检测、移除标记数据流中所嵌入的水印信息,甚至将水印消息复制到其他未标记数据流中,且已有流水印技术的水印容量有限.

本文的贡献主要体现在以下几个方面:

1) 提出基于包序重排的新型流水印技术.据目前所知,流水印载体局限于包载荷、流速率和包时间 3 种,本文采用包序作为流水印载体,并提出基于包序重排的新型流水印技术.

2) 将纠错码理论引入水印编码中.为达到一定的鲁棒性,已有流水印技术往往采用提高嵌入程度和增加冗余等手段,隐蔽性差,且水印容量小,本文将纠错码理论引入水印编码中,并用实验验证了该方式在提高流水印技术鲁棒性和水印容量时的有效性.

3) 引入概率调制思想.水印嵌入位置的选择影响到流水印的机密性及水印嵌入成功率等,然而现有流水印技术对此缺乏关注.本文引入概率调制思想,将包序重排比例控制在一定范围内,保证了流水印技术的隐蔽性.

1 相关工作

按照流水印载体的不同,当前网络流水印技术主要有基于包载荷、基于流速率和基于包时间的流水印技术 3 种.由于基于包载荷的流水印技术[15]与具体应用层协议(如 telnet,rlogin)有关,难以处理加密流量,适应范围较窄,且易被检测和过滤.相比而言,后两者可有效处理加密数据流,逐渐成为流水印技术研究的热点.下面主要对基于流速率和基于包时间的流水印技术进行分析.

基于流速率的流水印技术通过调制数据流的速率进而达到嵌入水印、追踪数据流的目的.Fu 等人[1]通过电磁干扰方式把一种可识别的标记信息嵌入到无线网络流量中以有效追踪数据流,但该方法难以抵御数字过滤(digital filtering)技术的干扰.通过干扰发送者流量并细微改变其发送速率,Yu 等人[2]提出一种基于直序扩频(direct sequence spread spectrum,简称 DSSS)的流追踪技术,但其只适合流量速率固定的情况,然而大多数匿名通信系统(如网络浏览、即时通信和远程登录等)所产生流量的速率却是不固定的.

基于包时间的流水印技术通过调制数据包的时间信息进而达到嵌入水印、追踪数据流的目的.为了有效应对时间扰乱,Wang 等人[16]通过操控包间隔到达时延(inter-packet delay,简称 IPD)提出一种基于水印的流关联技术进行跳板攻击溯源.Wang 等人[3]提出基于间隔重心的流水印(interval centroid based watermarking,简称 ICBW)技术以有效应对各种流变换,如丢包、流混杂(flow mixing)和流分割(flow splitting)等.为了有效抵抗包重组(repacketization)的干扰,Pyun 等人[8]提出基于时间间隔的流水印(interval-based watermarking,简称 IBW)技术.Houmansadr 等人[17]提出 RAINBOW 流水印技术,只需增大或减少极小的包间隔到达时延(该时延比先前流

水印技术所使用时延小得多)即可达到非常好的检测率.

然而,一方面,为了应对网络传输引入和攻击者主动添加的时间扰乱,进而达到较好的检测率,基于流速率和基于包时间的流水印技术必须向数据流中引入大的时延[17](RAINBOW 技术[17]虽然采用的时延相对较小,但为使水印信息顺利嵌入目标数据流中,该技术需引入非常大的初始时延),导致攻击者可发起多流攻击[18]、均方自相关攻击[19]和时间分析攻击[20]等来检测、移除标记数据流中所嵌入的水印信息,甚至将水印消息复制到其他未标记数据流中误导追踪.另一方面,这些流水印技术受其载体及调制方式所限,水印容量小.

为此,本文采用包序作为水印载体,旨在探讨新型流水印技术.

2 包序及包乱序基础

包序是指数据包报头中的序号信息,如 TCP[21]中的序号、IPsec 数据包的认证头(authentication header,简称 AH)[22]和封装安全载荷(encapsulating security payload,简称 ESP)[23]中的序号等.包乱序是指接收端收到数据包的顺序与发送端发送数据包的顺序不符.如果采用包序作为流水印载体,对包序进行调制进而嵌入水印信息的话,那么对于第三方而言,其表现形式就是包乱序.

目前,对于包乱序的研究集中于包乱序发生的可能性、影响因素、对网络稳定性和性能的影响[24]及如何降低或消除包乱序[25]等方面.另外,对于包乱序的高效或精确测量[26−28]也是当前的研究热点.

本节主要分析自然引入的包乱序和匿名通信系统等引入的包乱序,在此基础上,分析采用包序嵌入水印的可行性.

2.1 自然引入的包乱序

Bennett 等人[29]的研究结果表明,90%的网络会话都有数据包乱序情况发生,乱序比例在 0.1%~3%.已有研究成果[25−27,30,31]表明,造成包乱序的原因主要有高速路由器内置的并行性、数据包多经传输、负载均衡、链路层和 TCP 重传等.相对而言,前者影响更大,路由器为了达到更快的处理速度必然要增强并行处理能力.这些因素在现代数据网络中将长期存在,因此包乱序近期内难以彻底消除.于是,通过故意调整数据包顺序(控制在一定程度)来嵌入水印造成的包重排序会被视为自然行为.而且,引入更多的重排序数据包并不一定会降低整体会话和网络的性能[32].

2.2 匿名通信系统引入的包乱序

为躲避监管和提供匿名性,匿名通信系统和攻击者在包序上也会主动引入一些调整和变换,下面以匿名通信系统为例进行介绍.

在匿名通信研究领域,Mix[33]已经扩展为一种通用的匿名保护技术[34,35],人们基于此设计了大量匿名协议和系统(如 Onion Routing[36],Tor[37]和 Tarzan[38]等).Mix 网络通过改变到达信息的外在表现、移除其到达顺序来提供匿名性[34],其重要组成部分是 Mix 节点,通过加密来改变输入信息的表现形式,通过对输入接口的信息进行批处理和排序来隐藏到达信息的时间顺序,使得第三方无法确定输入消息和输出消息的对应关系,进而无法跟踪某条消息的传输路径.

虽然已有基于流速率和基于包时间的流水印技术不受加密的影响,然而,它们却依赖于包时延信息,因此难以从根本上应对 Mix 节点引入的时间扰乱,特别是在包乱序干扰存在的情况下,其检测率大为降低.

2.3 采用包序嵌入水印的可行性分析

如果可以使用包序作为流水印载体,那么不但可与数据包载荷和加密无关,对于网络时延和抖动也不那么敏感,鲁棒性会更强.

然而,流水印载体应可用于携带隐藏信息,如通过调制流速率载体使其呈现特定模式即可携带特定信息,同时嵌入水印信息后,载体的变化应不易被检测,如向包时间载体嵌入标记后,其表现形式与网络传输引入的时间扰乱类似.

对于包序而言,通过对数据包的不同顺序进行变换来代表不同的标记信息,即可嵌入标记信息,同时对包序

进行变换后,其表现形式为包乱序,而在互联网环境中,数据包乱序时常发生[39],而且由于处理代价过高等原因,包乱序在实际中往往被忽视[26,30,32,39],因此,采用包序作为流水印载体是可行的[40−42].

3 基于包序重排的新型流水印技术

为权衡鲁棒性(robustness)、隐蔽性(stealthiness)和容量(capacity)等需求,采取以下手段设计基于包序重排的新型流水印(packet reordering based flow watermarking,简称 PROFW)技术:

1) 引入检错和纠错机制提高流水印技术的错误抵抗能力和鲁棒性[43];
2) 为保证流水印的隐蔽性,引入概率调制思想[44,45],牺牲部分容量,将整体包序重排程度控制在正常范围内.

下面对 PROFW 技术的分层模型、水印信息编码与译码过程、数据流概率调制与解调过程等进行介绍.

3.1 PROFW分层模型

PROFW 技术采用分层模型,如图 1 所示,该模型由水印嵌入(watermark embedding)模块和水印检测(watermark detection)模块组成.

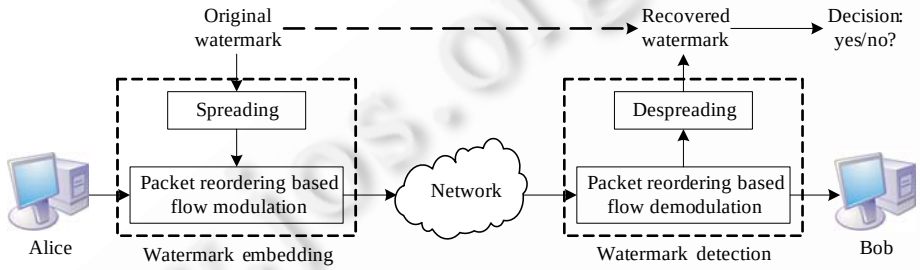

Fig.1 PROFW layered model

图 1 PROFW 分层模型

水印嵌入模块负责对原始水印进行编码并调制目标数据流.首先将原始水印信号 w 编码为特定个数(设为 k)数据包的不同排列,接着从数据流中选择适量的数据包组(使得调制后的标记数据流的整体乱序不超过特定阈值),通过缓存这些数据包组,并按照编码后的不同排列对其中的数据包顺序进行调整,以此将水印信息嵌入到目标数据流 F^u 中.然后标记数据流被送入网络(如匿名网络或跳板链等)中进行传输.

水印检测模块的功能是从受干扰后的标记数据流中恢复出其中所嵌入的水印信息.首先从接收或观察到的标记数据流 F^w 中选择得到相同的数据包组,从数据包组中获得数据包序号信息及到达顺序,在检错和纠错后译码恢复出水印信息 w'.如果恢复水印 w'与原始水印 w 相同,就认为主机 Alice 和主机 Bob 之间存在通信关系.

与传统的密码和数字水印体制类似,流水印技术也依赖于水印嵌入模块与水印检测模块之间的共享秘密.假设下面的参数是水印嵌入模块与水印检测模块所共享(如通过某种秘密途径协商或预先设定等):水印 w、包组中包数 k、所采用码字集合 M、编译码方式和乱序比例阈值.

3.2 水印信息编码和译码过程

水印信息编码的核心问题是从包组中选取具有检错和纠错能力的数据包排列用于代表水印信息.由于译码为编码的逆过程,下面主要介绍水印信息编码过程.

对流水印技术而言,鲁棒性是基本要求,即在面临强干扰时仍能保持一定的准确率.为了提高 PROFW 技术的鲁棒性,引入了检错和纠错机制.显然,对于长度为 k 的码字有 $k!$种排列方式,但为了实现检错和纠错,只有能够使用其中的部分码字(假定满足错误处理要求的码字集合为 M,码字个数为 m),其码率为 $r=\log_2 m/\log_2 k!$.

为了能够处理数据包移位错误和防止错误传播[42],根据汉明距离(Hamming distance)[46]建立映射关系,这里使用格雷码(Gray code)[47]进行编码,每个数据包移位对应单个汉明距离.示例见表 1,其中 $k=3,m=6$.

一般而言,为检测 N_e 个错误,码字间需要的最短距离(即最小码距)为 N_e+1,为纠正 N_e 个错误,需要的最小码

距为 $2N_e$+1.若想检测一个错误,那么所选码字之间的最小汉明距离为 2,所选码字应为{CW_1,CW_3,CW_5};若想纠正一个错误,那么最小汉明距离为 3,可使用的码字只有{CW_1,CW_4},可检测两个错误.

Table 1 An example of packet reordering encoded by Gray code
表 1 对包序采用格雷码编码示例

Codeword	Gray code
CW_1=〈1,2,3〉	00 00
CW_2=〈1,3,2〉	00 01
CW_3=〈3,1,2〉	10 01
CW_4=〈3,2,1〉	10 11
CW_5=〈2,3,1〉	10 10
CW_6=〈2,1,3〉	00 10

3.3 数据流概率调制与解调过程

为保证流水印的隐蔽性,引入概率调制思想,将 PROFW 技术对目标数据流的调制所导致的包乱序控制在一定范围内.

3.3.1 数据流概率调制

在对水印信息进行编码之后,对包序的调制相对简单.但为满足隐蔽性,必须将因嵌入水印所造成的包乱序控制在一定范围内.本文采用乱序密度(reorder density,简称 RD)[27,48]来度量标记数据流的乱序程度.

码字集合 M 中含有 m 个码字,为使传输带宽最大,码字应同等概率使用,则每码字的信息量是 $\log_2 m$ 比特,即带宽为 $\log_2 m$ 比特/码字或$(\log_2 m)/k$ 比特/包.为增强隐蔽性,应使标记数据流的乱序比例不与正常行为偏离太多而引起报警或被检测到.以牺牲部分带宽和容量为代价,PROFW 技术通过限制“乱序”码字的使用率来提高隐蔽性.假定 p_M 为编码水印消息时码字集 M 中码字的概率分布,那么每码字的信息量降为 $H(p_M)$比特.

为方便理解,以 k=m=2(码字分别为 CW_1=〈1,2〉和 CW_2=〈2,1〉)为例.假设正常情况下数据流的 RD 为 RD_{normal}=〈0.95,0.05〉,即 95%的数据包按顺序到达,只有 5%的数据包乱序.码字 CW_i 对 RD 的贡献不同,对于 CW_1,其归一化 RD 为 RD_1=〈1,0〉,CW_2 对应的 RD 为 RD_2=〈0.5,0.5〉,则有:

$$\begin{pmatrix} 1 & 0.5 \\ 0 & 0.5 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix}=\begin{pmatrix} 0.95 \\ 0.05 \end{pmatrix} \tag{1}$$

其中,p_i 表示码字 CW_i 的概率.求解可知,p_1=0.9,p_2=0.1,即 90%情况下需按原序发送,而逆序发送的比例不能超过 10%,否则就会引起异常.

一般地,需求解 $RD_M \times P_M = RD_1 RD_{normal}$,其中每列 RD_i 是 k 大小的向量,表示码字造成的 RD,P_M 是 m 个码字待求解的概率向量.

待求解获得各码字概率之后,包组选择函数依据该概率分布选择合适的数据包组.

3.3.2 数据流解调

数据流解调过程相对简单,水印检测模块采用预先分配或经由其他途径(与水印嵌入模块)协商的参数 k,M 和最大允许 RD(或其他类似度量方式,用于设定码字使用率),再依据包组选择函数选择得到与水印嵌入模块相同的数据包组,将接收或观察(嗅探)到的包组中数据包报头中的序号信息(如 TCP[21]中的序号、IPsec 数据包的认证头[22]和封装安全载荷[23]中的序号等)提取出来(无需对数据包进行缓存等复杂处理),依据其到达顺序将其放入到相对应的码字(包组)队列,每当一个码字中所有数据包序号都接收到之后,对其进行检错和纠错处理(在需要和可能的情况下),即可解调得到相应码字.

4 PROFW 技术水印容量分析

由于 PROFW 技术的隐蔽性由概率调制保证,而对于时间扰乱和包乱序等干扰的鲁棒性由纠错编码保证,下面主要分析 PROFW 技术的水印容量.

对于包含 k 个数据包的码字,可纠正的最大错误数为 $e_c \leqslant \lfloor (n-1)/2 \rfloor$(其中 $n \approx k\log_2 k$),可检测的最大错误

数为 $e_d \leqslant n-1$,最大值发生在最小码率 $r=1/n=1/k\log_2 k$ 时(即采用 $k!$个排列中的两个来代表水印信息).基于参数 k,m,p_M 和错误处理要求,PROFW 技术的水印容量为

$$C_{\text{PROFW}}=\frac{H(p_M)}{k}\text{(比特/包)} \tag{2}$$

若忽略隐蔽性要求,$H(p_M)$可用其最大值 $\log_2 m$ 代替,有:

$$C_{\text{PROFW}}=\frac{1}{k}\log_2 m\text{(比特/包)} \tag{3}$$

其中 m(对于线性分组码)为

$$m\leqslant\frac{2^n}{\sum_{i=0}^{d}\binom{n}{i}} \tag{4}$$

d 为码字间最小汉明距离(为达到特定检错、纠错能力).于是有 PROFW 技术的整体水印容量为

$$C_{\text{PROFW}}\leqslant\frac{1}{k}\log\left(\frac{2^{\lceil k\log_2 k\rceil}}{\sum_{i=0}^{d}\binom{\lceil k\log_2 k\rceil}{i}}\right)\text{(比特/包)} \tag{5}$$

从式(5)可以看出,随着最小码距 d 的增加,水印容量 C_{PROFW} 随之减小,而随着码字长度 k 的增加,水印容量 C_{PROFW} 随之增加.因此,PROFW 技术需要权衡参数码字长度 k 和最小码距 d.

5 实验结果与分析

为测试 PROFW 技术的有效性,搭建如图 2 所示的实验环境.

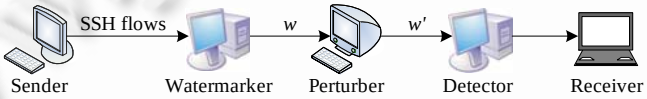

Fig.2 Experimental setup

图 2 实验环境

发送端与接收端之间使用 SSH 数据流进行通信;在水印嵌入器(watermarker)处,将水印 w 嵌入目标数据流中;标记数据流在流经干扰器(perturber)时,干扰器对其进行时间扰乱、包乱序等干扰;最后水印检测器(detector)从受干扰后的标记数据流中解调和译码恢复出水印 w',将其与原始水印 w 比对.

实验中采用互联网数据分析合作协会(Cooperative Association for Internet Data Analysis,简称 CAIDA)的 equinix-chicago 监视器于 2011 年 1 月捕获的数据集[49]作为测试数据集,从中提取 SSH 数据流作为发送端的数据流.实验中选取 80 条 SSH 数据流,每条数据流中至少包含 5 000 个数据包.默认情况下,干扰器对标记数据流进行包乱序干扰的比例为 13%(实际环境中的乱序比例一般在 0.1%~3%之间[29]).

5.1 严重干扰下的错误率

经网络传输与干扰后,数据包实际到达顺序与码字中的预期顺序不一致就会产生错误.图 3 显示了 PROFW 技术在码字长度变化时的错误率情况.测试结果显示,同一码字中,错误率随着错误数量的增加而急剧下降,表明同一码字中产生多个错误的概率很小,因此,PROFW 技术的鲁棒性较好,具有较强的错误抵抗能力.

5.2 降低容量时的错误率

为测试 PROFW 技术在牺牲水印容量情况下的错误率,本次实验只使用 $k!$个码字中的两个码字来嵌入水印.图 4 显示了在不同错误数量情况下,随着码字长度 k 的增加,PROFW 技术的累计错误率情况.

由图 4 可知,在特定码字长度 k 的情况下,错误数越多,累计错误率越高,即纠错能力越强,正确率越高.如在具

备纠正单个错误($k\geqslant 3$ 时即可)的能力时,码字被成功接收的概率高达 98.2%.

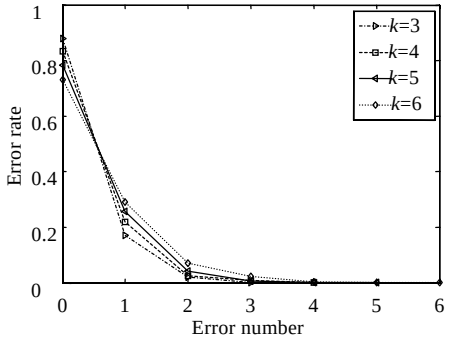

Fig.3 Error rate of PROFW

图 3 PROFW 技术的错误率

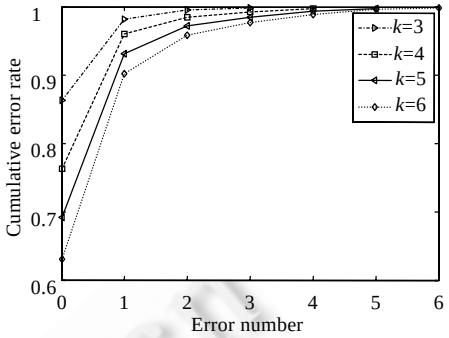

Fig.4 Cumulative error rate under different codeword sizes

图 4 不同码字长度下的累计错误率

5.3 特定错误处理能力下的水印容量

当码字具备特定的错误处理能力(以最小码距 d 表示)时,测试 PROFW 技术的水印容量,测试结果如图 5 所示.

由图 5 可知,在特定错误处理能力下,随着码字长度 k 的增加,每个数据包可承载的字节数也随之增加.同时,在特定码字长度情况下,随着最小码距 d 的增加,水印容量随之减小,这与式(5)的分析结果高度吻合.如在 k=6 时,随着最小码距 d 的增加,水印容量从 1.73 比特/包~0.65 比特/包,在 d=2 时,随着码字长度 k 的增加,水印容量从 0.34 比特/包~1.82 比特/包(这是其他流水印技术难以达到的).

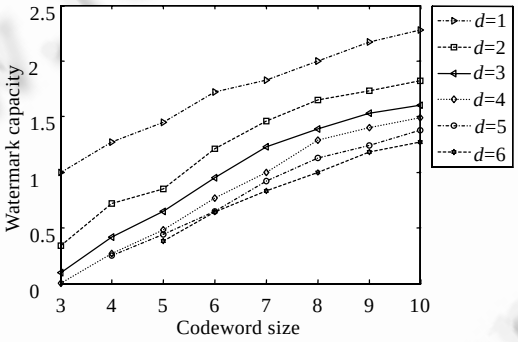

Fig.5 Watermark capacity under different codeword sizes

图 5 不同码字长度下的水印容量

5.4 与其他流水印技术的检测率对比

本次实验测试 PROFW 技术与其他典型流水印技术(DSSS[2],IBW[8],ICBW[3]和 RAINBOW[17])在使用特定数据包数量时的检测率对比,首先,在不进行包乱序干扰(但存在 2 000ms 的随机时间扰乱)时测试各流水印技术的检测率,各典型流水印技术均采用其典型配置参数,测试结果如图 6 所示.可以看到,PROFW 技术的检测率随着数据包使用量的增加而增加,且优于其他典型流水印技术,这是因为 PROFW 技术采用包序作为水印载体,受时间扰乱的影响较小.

然后,在干扰器处增添包乱序干扰(比例为 10%),各流水印技术的检测率测试结果如图 7 所示.可以看到,典型流水印技术的检测率下降严重,难以有效应对包乱序干扰,而 PROFW 技术因具备纠错能力,所以能够保持良好的检测率,说明 PROFW 技术在应对时间扰乱和包乱序干扰时的鲁棒性更强.

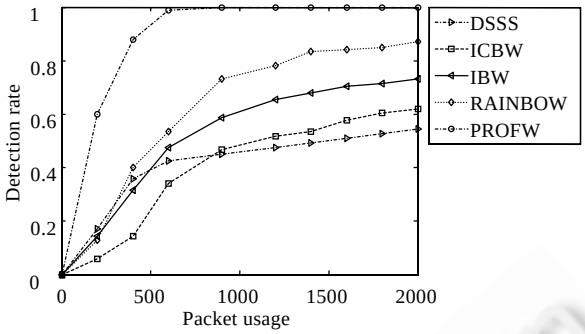

Fig.6 Detection rate comparison without packet reordering interference
图 6 无包乱序情况下的检测率对比

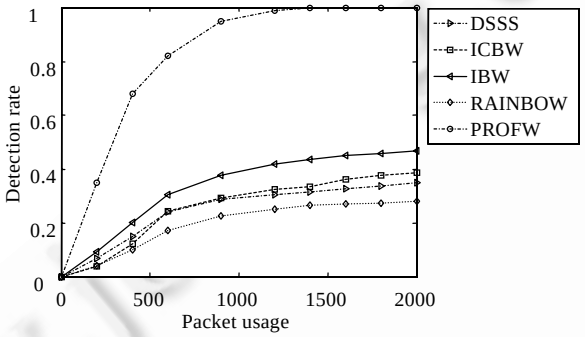

Fig.7 Detection rate comparison with packet reordering interference
图 7 包乱序情况下的检测率对比

6 结束语

当前流水印载体局限于包载荷、流速率和包时间这 3 种,本文采用包序作为流水印载体,并引入纠错码理论和概率调制思想,提出基于包序重排的新型流水印技术 PROFW.测试结果表明,PROFW 技术在保证隐蔽性的前提下,对于时间扰乱和严重包乱序具有较强的鲁棒性.与当前典型流水印技术相比,该技术不但在应对时间扰乱和包乱序时的鲁棒性更强,而且提高了水印容量.

下一步工作将研究该技术面临其他流变形情况(如向标记数据流中增加垃圾包或从标记数据流中移除数据包等)时的鲁棒性;探讨该技术对其他度量方式的隐蔽性,如乱序缓存占用密度(reorder buffer-occupancy density,简称 RBD)[28];由于数据包乱序在不同的网络和路径上存在很大的差异,因此 PROFW 技术的自适应性(数据包选择、嵌入成功率等)也是下一步的研究工作.

References:

[1] Fu XW, Zhu Y, Graham B, Bettati R, Zhao W. On flow marking attacks in wireless anonymous communication networks. In: Proc. of the 25th IEEE Int'l Conf. on Distributed Computing Systems (ICDCS). Columbus: IEEE Computer Society, 2005. 493−503.

[2] Yu W, Fu XW, Graham S, Xuan D, Zhao W. DSSS-Based flow marking technique for invisible traceback. In: Proc. of the 2007 IEEE Symp. on Security and Privacy (SP). Oakland: IEEE Computer Society, 2007. 7−21.

[3] Wang XY, Chen SP, Jajodia S. Network flow watermarking attack on low-latency anonymous communication systems. In: Proc. of the 2007 IEEE Symp. on Security and Privacy (SP). Oakland: IEEE Computer Society, 2007. 116−130.

[4] Wang XY, Chen SP, Jajodia S. Tracking anonymous peer-to-peer VoIP calls on the Internet. In: Proc. of the 12th ACM Conf. on Computer and Communications Security (CCS). Alexandra: ACM Press, 2005. 81−91.

[5] Peng P, Ning P, Reeves DS, Wang XY. Active timing-based correlation of perturbed traffic flows with chaff packets. In: Proc. of the 25th IEEE Int'l Conf. on Distributed Computing Systems Workshops (ICDCSW). Columbus: IEEE Computer Society, 2005. 107−113.

[6] Zhang LF, Persaud AG, Johnson A, Guan Y. Detection of stepping stone attacks under delay and chaff perturbations. In: Proc. of the 25th IEEE Int'l Performance Computing and Communications Conf. (IPCCC). Phoenix: IEEE Press, 2006. 247−256.

[7] Park YH, Reeves DS. Adaptive timing-based active watermarking for attack attribution through stepping stones. In: Proc. of the 2nd Int'l Workshop on Security in Distributed Computing Systems. Washington: IEEE Computer Society, 2007. 107−113.

[8] Pyun YJ, Park YH, Wang XY, Reeves DS, Ning P. Tracing traffic through intermediate hosts that repacketize flows. In: Proc. of the 26th IEEE Int'l Conf. on Computer Communications (Infocom). Anchorage: IEEE Press, 2007. 634−642.

[9] Pan Z, Peng H, Long XZ, Zhang CL, Wu Y. A watermarking-based host correlation detection scheme. In: Proc. of the 2009 Int'l Conf. on Management of e-Commerce and e-Government. Nanchang: IEEE Computer Society, 2009. 493−497.

[10] Ramsbrock D, Wang XY, Jiang XZ. A first step toward live botmaster traceback. In: Proc. of the 11th Int'l Symp. on Recent Advances in Intrusion Detection (RAID). Boston: Springer-Verlag, 2008. 59−77.

[11] Donoho DL, Flesia AG, Shankar U, Paxson V, Coit J, Stanford S. Multiscale stepping-stone detection: detecting pairs of jittered interactive streams by exploiting maximum tolerable delay. In: Proc. of the 5th Int'l Symp. on Recent Advances in Intrusion Detection (RAID). Zurich: Springer-Verlag, 2002. 17−35.

[12] Blum A, Song D, Venkataraman S. Detection of interactive stepping stones: Algorithms and confidence bounds. In: Proc. of the 7th Int'l Symp. on Recent Advances in Intrusion Detection (RAID). Sophia Antipolis: Springer-Verlag, 2004. 258−277.

[13] He T, Tong L. Detecting encrypted stepping-stone connections. IEEE Trans. on Signal Processing, 2007,55(5):1612−1623.

[14] Zhu Y, Fu XW, Gramham B, Bettati R, Zhao W. Correlation-Based traffic analysis attacks on anonymity networks. IEEE Trans. on Parallel and Distributed Systems, 2010,21(7):954−967.

[15] Wang XY, Reeves DS, Wu SF, Yuill J. Sleepy watermark tracing: An active network-based intrusion response framework. In: Proc. of the 16th Int'l Conf. on Information Security (IFIP/Sec). Paris: Kluwer Academic Publishers, 2001. 369−384.

[16] Wang XY, Reeves DS. Robust correlation of encrypted attack traffic through stepping stones by manipulation of interpacket delays. In: Proc. of the 10th ACM Conf. on Computer and Communications Security (CCS). Washington: ACM Press, 2003. 20−29.

[17] Houmansadr A, Kiyavash N, Borisov N. RAINBOW: A robust and invisible non-blind watermark for network flows. In: Proc. of the 16th Annual Network & Distributed System Security Symp. (NDSS). San Diego: The Internet Society, 2009. 224−236.

[18] Kiyavash N, Houmansadr A, Borisov N. Multi-Flow attacks against network flow watermarking schemes. In: Proc. of the 17th USENIX Security. San Jose: USENIX Association, 2008. 307−320.

[19] Jia WJ, Tso FP, Ling Z, Fu XW, Xuan D, Yu W. Blind detection of spread spectrum flow watermarks. In: Proc. of the 28th IEEE Int'l Conf. on Computer Communications (Infocom). Rio de Janeiro: IEEE Computer Society, 2009. 2195−2203.

[20] Peng P, Ning P, Reeves DS. On the secrecy of timing-based active watermarking trace-back techniques. In: Proc. of the 2006 IEEE Symp. on Security and Privacy (SP). Berkeley: IEEE Computer Society, 2006. 334−349.

[21] Postel J. Transmission control protocol. Request for Comments: 793, 1981. http://www.rfc-editor.org/rfc/pdfrfc/rfc793.txt.pdf

[22] Kent S. IP authentication header. Request for Comments: 4302, 2005. http://www.rfc-editor.org/rfc/pdfrfc/rfc4302.txt.pdf

[23] Kent S. IP encapsulating security payload (esp). Request for Comments: 4303, 2005. http://www.rfc-editor.org/rfc/pdfrfc/rfc4303.txt.pdf

[24] Laor M, Gendel L. The effect of packet reordering in a backbone link on application throughout. IEEE Network, 2002,16(5):28−36.

[25] Leung KC, Li VOK, Yang DQ. An overview of packet reordering in transmission control protocol (tcp): Problems, solutions, and challenges. IEEE Trans. on Parallel and Distributed Systems, 2007,18(4):522−535.

[26] Piratla NM, Jayasumana AP. Metrics for packet reordering—A comparative analysis. Int'l Journal of Communication Systems, 2007,21(1):99−113.

[27] Piratla NM, Jayasumana AP, Bare AA. Reorder density (rd): A formal, comprehensive metric for packet reordering. In: Proc. of the 2005 IFIP Networking Conf. on Networking. Waterloo: Springer-Verlag, 2005. 233−272.

[28] Piratla NM, Jayasumana AP, Bare AA, Banka T. Reorder buffer-occupancy density and its application for measurement and evaluation of packet reordering. Computer Communications, 2007,30(9):1980−1993.

[29] Bennett JCR, Partridge C, Shectman N. Packet reordering is not pathological network behavior. IEEE/ACM Trans. on Networking, 1999,7(6):789−798.

[30] Paxson V. End-to-End routing behavior in the Internet. IEEE/ACM Trans. on Networking, 1997,5(5):601−615.

[31] Piratla NM, Jayasumana AP. Reordering of packets due to multipath forwarding—An analysis. In: Proc. of the IEEE Int'l Conf. on Communications (ICC). Istanbul: IEEE Press, 2006. 829−834.

[32] Arthur CM, Lehane A, Harle D. Keeping order: determining the effect of tcp packet reordering. In: Proc. of the 3rd Int'l Conf. on Networking and Services (ICNS). Washington: IEEE Computer Society, 2007. 116−122.

[33] Chaum D. Untraceable electronic mail, return addresses, and digital pseudonyms. Communication of the ACM, 1981,24(1):84−88.

[34] Sampigethaya K, Poovendran R. A survey on mix networks and their secure applications. Proc. of the IEEE, 2006,94(12): 2142−2181.

[35] Edman M, Yener B. On anonymity in an electronic society: A survey of anonymous communication systems. ACM Computing Surveys, 2009,42(1):132−166.

[36] Goldschlag D, Reed M, Syverson P. Onion routing for anonymous and private Internet connections. Communications of the ACM, 1999,42(2):39−41.

[37] Dingledine R, Mathewson N, Syverson P. Tor: The second-generation onion router. In: Proc. of the 13th USENIX Security Symp. San Diego: USENIX Association, 2004. 303−320.

[38] Freedman MJ, Morris R. Tarzan: A peer-to-peer anonymizing network layer. In: Proc. of the 9th ACM Conf. on Computer and Communications Security (CCS). Washington: ACM Press, 2002. 303−320.

[39] Bellardo J, Savage S. Measuring packet reordering. In: Proc. of the 2nd ACM SIGCOMM Workshop on Internet Measurment (IMW). Marseille: ACM Press, 2002.

[40] Estévez-Tapiador JM, Castro JCH, Alcaide A, Ribagorda A. On the distinguishability of distance-bounded permutations in ordered channels. IEEE Trans. on Information Forensics and Security, 2008,3(2):166−172.

[41] Chakinala RC, Kumarasubramanian A, Manokaran R, Noubir G, Pandu RC, Sundaram R. Steganographic communication in ordered channels. In: Proc. of the 8th Information Hiding Workshop. Old Town Alexandria: Springer-Verlag, 2006. 42−57.

[42] Khan B, Paduch J, Levy J. Superimposing permutational covert channels onto reliable stream protocols. In: Proc. of the 3rd Int'l Conf. on Malicious and Unwanted Software (MALWARE). Fairfax: IEEE Computer Society, 2008. 49−56.

[43] Medeni MBO, Souidi EM. A novel steganographic protocol from error-correcting codes. Journal of Information Hiding and Multimedia Signal Processing, 2010,1(4):337−343.

[44] Fridrich J, Goljan M. Digital image steganography using stochastic modulation. In: Proc. of the SPIE 5020: Security and Watermarking of Multimedia Contents V. Santa Clara, 2003. 191−201.

[45] He JH, Huang JW. Steganalysis of stochastic modulation steganography. Science in China (Series F Information Sciences), 2006,49(3):273−285.

[46] Hamming RW. Error-Detecting and error-correcting codes. Bell System Technical Journal, 1950,29(2):147−160.

[47] Savage C. A survey of combinatorial gray code. SIAM Review, 1997,39(4):605−629.

[48] Banka T, Bare AA, Jayasumana AP. Metrics for degree of reordering in packet sequences. In: Proc. of the 27th Annual IEEE Conf. on Local Computer Networks (LCN). Los Alamitos: IEEE Computer Society, 2002. 333−342.

[49] Claffy K, Andersen D, Hick P. The caida anonymized 2011 Internet traces. 2011. http://www.caida.org/data/passive/passive_2011_dataset.xml

张连成(1982－),男,河南商丘人,博士,CCF 学生会员,主要研究领域为流量分析,网络安全.

王振兴(1959－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为流量分析,网络与信息安全.

徐静(1973－),女,博士生,主要研究领域为流量分析,匿名通信,网络安全.