主动网络流水印技术研究进展 *

金 华，王 成

(江苏大学 计算机科学与通信工程学院, 江苏 镇江 212013)

摘 要：随着互联网应用深入到社会的各个方面，网络安全与隐私保护形势日益严峻。匿名网络技术能较好地隐藏身份等个人隐私信息，但又易于被攻击者所利用，为其不法行为增加掩护，并阻碍对其监控与追踪。主动网络流水印技术是解决非法入侵追踪的主要方法。首先简要介绍主动网络流水印技术的应用场景；其次描述了水印基本框架，对典型的主动网络流水印技术进行了归纳和梳理，对各自的实现原理进行了分析比较，给出了最新研究进展；接着讨论了现有针对网络流水印的攻击方法，并从隐蔽性、鲁棒性、抗攻击性等六个方面对水印典型算法的性能进行了比较；最后总结并指明下一步的研究方向。

关键词：网络隐通道；流水印；流量分析；匿名通信；入侵追踪

中图分类号：TP391 doi: 10.19734/j.issn.1001-3695.2019.01.0003

Progress in research on active network flow watermark

Jin Hua, Wang Cheng

(*School of Computer Science & Communication Engineering*, *Jiangsu University*, *Zhenjiang Jiangsu* 212013, *China*)

Abstract: With the Internet application moving into societies increasingly, network security and privacy preservation are becoming a growing threat. Anonymous communication technology are better able to hide personal privacy information, but it is easy to be exploited by attackers, such as covering their illegal behavior and impeding to be monitored and tracked. Active network flow watermark is the main method to realize illegal intrusion traceback. The main scenes of network flow watermark were summed up. The basic framework of watermarking was described, and the typical active network flow watermark technologies were summarized. The realization principle of each method were analyzed in details, and the recent progress of active network flow watermark was presented. Then, the existing attack methods of network flow watermark were discussed, and the performance of watermark algorithms were compared on six aspects, such as concealment, robustness, anti-attack and so on. Finally, the future research of active network flow watermark was proposed.

Key words: network covert channel; flow watermark; traffic analysis; anonymous communication; intrusion traceback.

随着信息化时代的到来，网络安全越来越成为人们关注的焦点，网络入侵者为了自身利益，通过各种攻击手段威胁着网络安全，非法侵占他人资源。通常分为以下三种攻击场景[1]：a）基于网络的攻击，攻击者利用自己设计的跳板节点来窃听或修改与受害者之间的通信数据，利用僵尸网络（如 DDoS、垃圾邮件和钓鱼网络等）来获取浏览者的信息；b）安全和隐私攻击，攻击者收集和分析用户可能访问的网页特征，窃取 URL 信息，又如 VoIP 安全威胁，威胁主要来自于病毒、木马和黑客对数据网络的各种攻击，包括非法接听、话费诈欺等；c）匿名攻击，攻击者通过匿名通信系统（如 TOR、Mix、Crowds 等）传播暴力、毒品、色情等不良信息。

针对以上问题，早期的解决方式主要采用被动网络流分析法。研究表明，通过检查和被动分析可以收集有用的信息。例如，通过分析流分组的统计特征（分组长度、分组间隙、分组方向等）以及利用先进的机器学习算法，可以在给定的一组选择中定义应用协议的类型，预测用户的位置[2]，检测异常流量[3]，区分恶意流量[4]。不过，被动网络流分析也会被攻击者利用，用来推断关于用户通信的私人或敏感信息，这种类型的机密性违规示例有识别已加密和认证了的已下载网页；识别加密 VoIP 通信中的会话；获得加密 VoIP 会话的部分副本。总体来看，被动网络流分析有三个主要缺点：a）它需要使用复杂的机器学习算法，通常不能在可扩展性和准确性之间达到最佳平衡；b）提前需要大量的样本流来训练机器学习算法；c）网络行为易受通信干扰的影响以及攻击方对流量的恶意操纵。何高峰等人[5]提出了 TOR 匿名通信流量在线识别方法（基于 TLS 指纹和基于报文长度分布的识别方法），不过此方法只对于 TOR 系统效果显著，应用范围单一。

为了解决上述问题，研究者已经根据数字水印的思想：在数字内容中嵌入专有信息(即水印，永久性嵌入数据中的识别码，并且在任何解密过程之后仍然存在于数据内)的方法。在隐通道的基础上提出了一种主动网络流水印技术（active network flow watermarking，ANFW）。ANFW 通过改变发送端中产生的网络流指定特征来嵌入水印，然后在接收端检测对应流中是否存在水印，从而判断出发送端和接收端是否存在流关联。这种主动流水印技术比传统的被动流水印技术更具适应性，它完全可以存在于匿名通信网络中或是其他网络环境下，这对于检测非法通信效果显著。因此，该技术近年来逐渐成为网络安全研究领域的热点。

1 水印框架与典型 ANFW 技术

1.1 水印架构

在给出网络流水印的一般框架之前，首先引入隐通道的概念。隐通道已经讨论了几十年，隐通道被定义为“以违反系统安全策略的方式传输信息的过程中可以利用的任何通信

收稿日期：2019-01-05；**修回日期：**2019-02-23 **基金项目：**国家自然科学基金资助项目（61672269，61300228，31471646）；江苏省科技成果转化项目（BA2015161）

作者简介：金华（1977-），男，江苏海安人，副教授，硕导，主要研究方向为网络信息安全、隐私保护；王成（1993-），男，江苏兴化人，硕士研究生，主要研究方向为网络信息安全（1185778813@qq.com）.

渠道”。Zander 等人[6]进一步缩小并确定其为网络协议中的信息隐藏的子类“网络隐通道”，并给出了四种可能的通信场景，如图 1 所示。具体场景取决于隐通道传送的发送者和接收者是否也是合法通信的接收者和发送者。在第二种场景下，合法通信的发送方和接收方也是隐蔽通信的发送方和接收方。第三种场景下，合法和隐蔽的发送者相同，而隐蔽与合法接收者不同，前者不知道他的通信被第三方利用为隐通道。相比之下第四种场景下，合法和隐蔽的接收者是相同的，而两个发送者是不同的。第一种场景下，两个隐通道端点与合法端点完全不同，这也是主动网络流水印的基本框架。

网络流水印技术通过水印机和水印检测器两个主要部分实现。两部分的位置是基于所追求的目标和期望观察的目标流而选择设计的。水印机负责将信息转换为具有某些特定属性的水印码，并将其嵌入到目标流中；相比之下，水印检测器观察网络中特定点处的通信流，并分析流量的特征，检测水印流并解码水印，从而获得嵌入其中的水印信息。

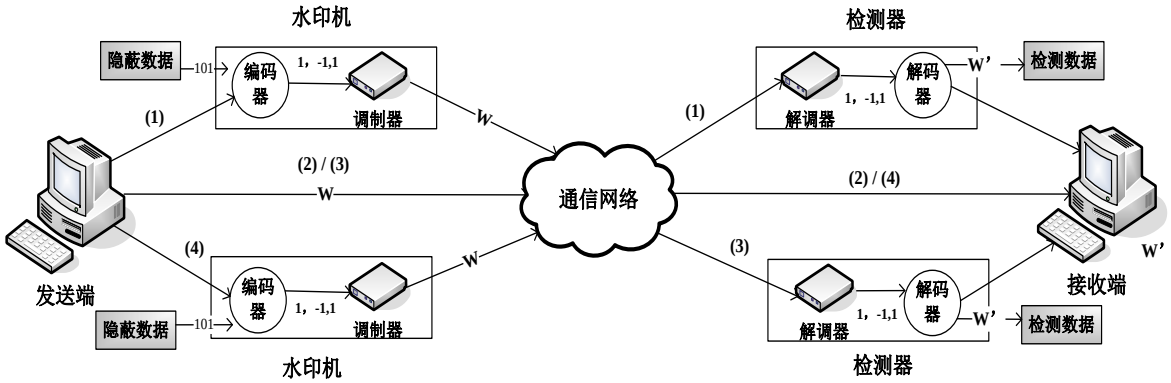

图 1　隐通道基本场景（1、2、3、4）及水印架构

Fig. 1　Basic scene and watermark architecture of covert channels(1,2,3,4)

1.2　典型 ANFW 技术

ANFW 技术主要是选择网络流中某些独立于数据包内容的特征来作为水印参数，被选取的网络流特征，称为水印载体。当网络流通过检测器时，通过解码来进行水印识别。根据水印载体的选择，ANFW 技术主要包括基于流速、基于分组延迟和基于时隙特征等[7]。图 2 给出了主动网络流水印技术的发展路线。从图中可以看出 ANFW 一直是网络安全领域研究的热点。

1.2.1 基于流速的 ANFW

基于流速的 ANFW 是选取整个流持续时间内的不同时间段，在不影响整个流传输的情况下，采取相应策略改变对应时间段内的流速率模式，将流速的不同变化表示为水印信息位 W_i（1 或 0），从而完成水印嵌入工作。

在基于流速的 ANFW 技术中，其典型方法是 Yu 等人[8]提出的方案：将无线通信系统使用的 DSSS(direct sequence spread spectrum)机制加入到了 ANFW 中，该方案被称为 DSSS-W。DSSS-W 在发送端将原始信号 W_i 按照对应 PN 码扩频为 W_i^D,接着在扩频期间按照一定的调节幅度对流速率进行略微改变，从而完成整个水印的嵌入。

然后，在接收端利用高通滤波器对接收到的速率滤除直流分量，即去除流的原始平均流速；接着，加上 PN 码再利用低通滤波器滤除噪声；最后，应用简单的决策规则恢复出每一个水印信息 W_i。

1.2.2 基于分组延迟的 ANFW

基于分组的网络流水印选择分组间延迟（inter-packet delay， IPD）作为水印载体，以获得更好的相关分析结果。IPD 是到达和出发时间之间的数据包的间隔。本文可以调整一些所选 IPD 或 IPD 的平均值大小来嵌入水印信息位 W_i。

为了解决攻击者的时间扰动,Wang 等人[9]提出了一种新型水印方案。通过随机选择两个包的数据流并计算其 IPD，然后使用式(1)将其变成一个新值 IPD^w。

$$IPD^w = EM(IPD,W,s) = \left[q\left(IPD + \frac{s}{2}, s \right) + \Delta \right] s,$$

$$\Delta = \left(W_i - \left(\mathrm{q}\left(\mathrm{IPD} + \frac{\mathrm{s}}{2}, \mathrm{s} \right) \bmod 2 \right) + 2 \right) mod\, 2 \qquad (1)$$

其中：s 为量化步长。之后，即可在发送端完成水印嵌入。

最后，在接收端利用式(2)破译水印相关网络流：

$$DE(IPD^w, s) = q(IPD^w, s) \bmod 2 \qquad (2)$$

其中：q（x,y）= round（x,y）表示取与 x 最近的整数。另外，方案指出使用平均 IPD 可以提高对攻击者时间扰动的鲁棒性。

Wang 等人方法的优点是它可以在短流中嵌入更多水印信息位。但是这种水印相关方法对于非独立随机延迟并不是很强大，并且当它们计算 IPDs 的平均值时，它们需要一个数据包缓冲来存储流的某些数据包，这将增加数据包的延迟，使其不能用于跟踪实时流。

为了解决这一问题,Wang 等人[10]在一篇研究对象为追踪匿名对等 VoIP 呼叫的文章中提出另一个网络流水印技术。该技术并没有量化数据流的 IPD,而是调整增量参数，以改变一组归一化 IPD 差异的平均值。因此，它可以根据呈现水印位的变化嵌入水印。

然而,通常基于 IPD 的水印方案不够安全,容易受到狡猾攻击者的攻击。Peng 等人[11]提出了一种基于相邻垫脚石之间的 IPD 分析攻击思想，并给出了一种推断水印参数的算法。为了对以上攻击取得良好的鲁棒性，研究人员使用非盲水印来分析流相关。Houmansadr 等人[12]提出了一种名为“RAINBOW” 非盲水印方案。

图 3 描述了 RAINBOW 网络流水印方案的模型。在水印框架的基础上增加了 IPD 数据库，用于存储 IPD。该方法的主要特点是：当在接收端检测水印时，它们需要流的起始 IPD 值。为了有效消除盲流中丢包和包重组的影响， RAINBOW 使用低于普通水印技术的延迟量，并保持其不可见性，以防止被攻击者检测并删除。不过，任何事情都有两面性,因为使用的非盲水印,此方案必须使用一个数据库来记录 IPD, 所以与其他盲水印方案相比，需要逐一匹配 IPD，具有较高的时间复杂度[13]。

另外，王昌达等人[14]在 2015 年提出了基于包交叉分组间隔质心双盲流水印技术，该方法通过将传输时间影响分散化，抵抗数据包在传输中遭遇的抖动干扰问题；通过约束数据包的传输时间分布来提高水印的抗检测性；通过利用卷积码将水印序列扩展后再传输，实现流水印的有限自纠错；通过基于滑窗的算法动态来判定数据包分组边界，实现流水印的盲追踪，并可抵御数据包在传输过程中可能出现的丢失或

合并影响。通过实验表明了此方法可在较高阈值范围情况下准确检测出水印的存在性，抵御一般性的第三方试探性检测，具有较强的鲁棒性与隐蔽性。

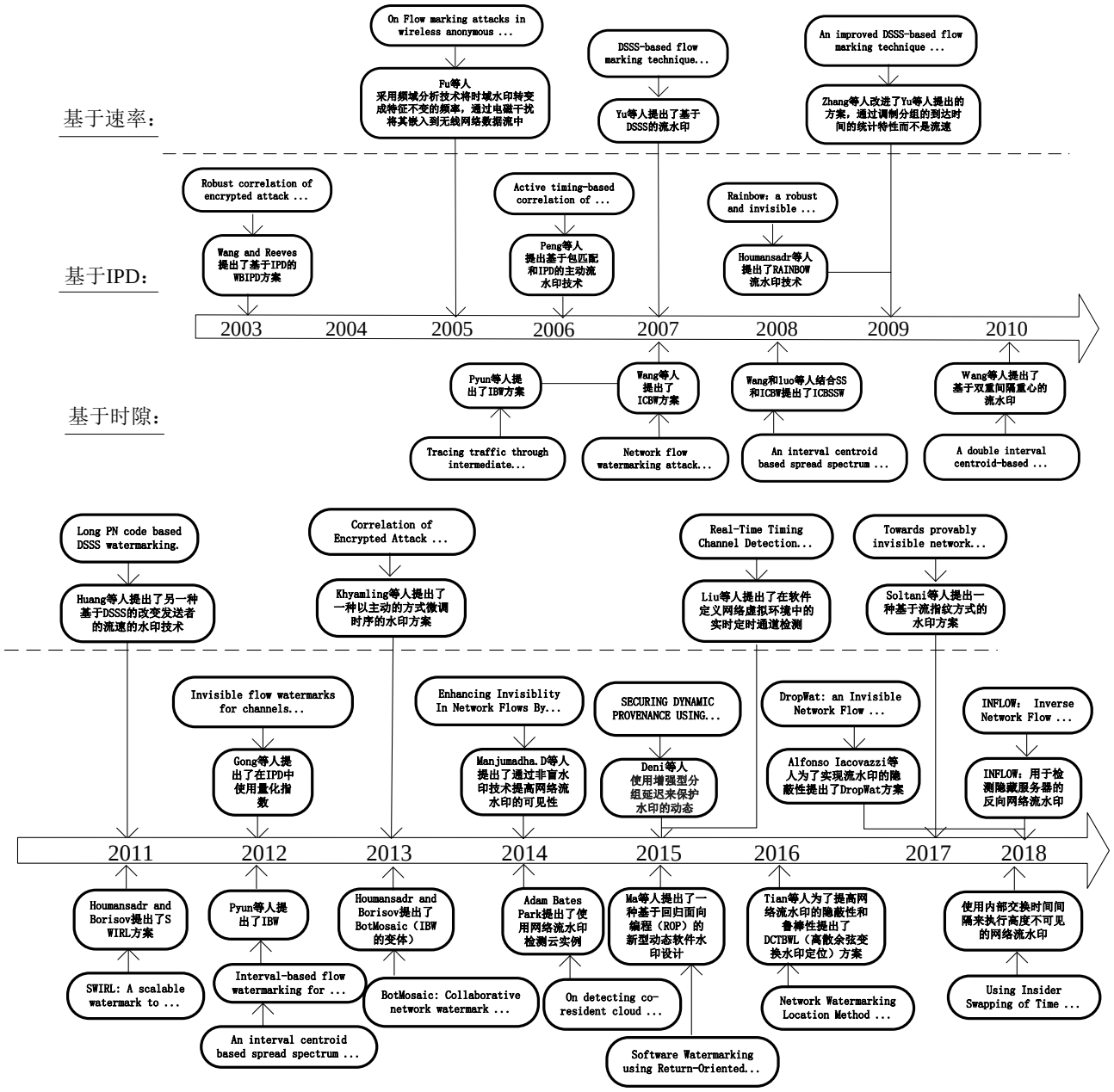

图 2　主动网络流水印技术的发展路线

Fig. 2　Development route of active network flow watermarking technology

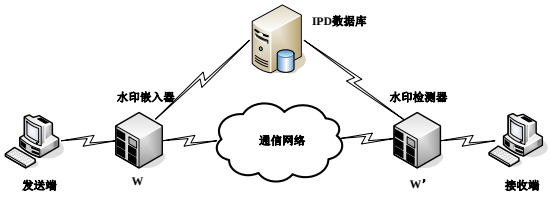

图 3　RAINBOW”网络流水印模型

Fig. 3　Network flow watermarking model:RAINBOW

Iacovazzi 等人[15]在 2018 年提出了一种用于数据泄露攻击的不可见的网络流水印技术 DropWat。DropWat 基于将一个水印注入到一个全新的范例。其中算法的基本思想是模拟自然网络行为删除流的几个选定分组，以便改变分组间延迟。该方法的优点是分组丢失很难被识别；自然丢失与有意向的丢失不易区分；嵌入此算法的水印是不可见的。

此外，Lacovazzi 等人[16]2018 年又提出了一种新的用于检测隐藏/目标服务器的反向网络流水印 INFLOW。基于匿名通信系统 TOR，利用 TCP 拥塞机制对 TOR 网络流的影响，INFLOW 在来自隐藏/目标服务器的接收端以短时间间隔丢弃数据包，丢包会影响 TOR 流控制，并导致在隐藏/目标服务器端观察到的流中存在时间间隔。通过控制通信边缘并检测水印间隙，INFLOW 能够检测隐藏/目标服务器。通过在真实 TOR 网络上的实验，结果显示真实的检测率在 90%~98%内。

1.2.3 基于时隙的 ANFW

基于 IPD 的水印技术被证明是跟踪匿名通信系统流量的有效方法，它对于攻击者涉及的有限数量的扰动具有较高的鲁棒性。然而 Pyun 等人[17]发现，当网络流在某些“垫脚石”上变换时，该技术不够健壮，如改变通信流的分组数，这些转换破坏了基于 IPD 的水印所需数据包的同步性。因此，他们提出了一种针对采用间隔作为水印载体的定时扰动和重新分组的新方法，该方法也称为基于间隔的水印(interval-based watermark, IBW)，将每个流的持续时间同步切割成固定长度的间隔（也称时隙），调整分组定时以在一定间隔内操纵分组内数据包的数量，嵌入水印 W_i。

人们总认为，匿名通信系统可以通过转换流量，如流填充、添加虚假数据包、流混合、流分解和流程合并等方式实现良好的匿名性。然而 Wang 等人[18]提出了一种时隙质心水印机制（interval centroid based watermarking，ICBW）。它通过在两个间隔中调整某些分组的时间偏移来嵌入水印位，即使是上述方式实现的匿名通信系统，只要数据流足够长，就能够识别出特定水印流。此方案首先给定持续时间 T_f，选择偏移量 σ>0（起始点），分割为 2n 个时隙 I_1，I_2,…，I_i，每个时隙持续时间为 T，第 i 个时隙 I_i 包含 n_i 个数据包。此时，选择平衡的特征如式(3)所示。

$$x_i = 1/n_i \sum_{h=1}^{n_i} [(t_h - \sigma) mod T] \tag{3}$$

其中：$i \in (1,2,\ldots,2n)$，t_h 是时隙 I_i 的第 h 个包的到达时间。

接着，将 2n 个时隙 I_i 随机分为数量相等的两组，以相同概率从每组中各取 r 个时隙，并计算此 r 个时隙的质心 A_i 和 B_i。由于 $Y_i = A_i - B_i$ 服从对称轴为 0 的均匀分布，所以可给 A_i 或 B_i 施加一个偏移量 a(0<a<T)，通过改变 Y_i 的对称轴完成水印位 W_i 的嵌入。

SWIRL(scalable watermark that is invisible and resilient to packet losses)[19]采用了一种时隙特征变换的流水印方法。选择的标记间隔被分成四个子间隔，每个子间隔被分成三个时隙。在每个子间隔中选择时隙，然后每个分组被延迟，使得所选择的时隙也被延迟。由于选择过程由随机水印参数控制，标记间隔中的模式是不同的。简单来说，其模式是根据被标记流的特征来选择，所以每个流都以不同的模式标记。SWIRL 还引入了小小的延迟，使其能够在现实中使用，并且被证明是一种有效的防御方式，可以抵御 MFA 攻击、TOR 拥塞攻击等。但在某些情况下，嵌入在 SWIRL 中的水印会被恶意攻击者检测到，攻击者能够通过传入特定的合法流让 SWIRL 重新嵌入水印信息，使得水印信息转移至目标流，再通过模式对比方式了解 SWIRL 嵌入水印信息的情况[20]。

Liu 等人[21]在 2018 年基于时隙提出了一种使用内部交换算法实现高不可见的网络流水印方案，提出了一种自适应的质心量化框架，使用近似正交的序列集来确定质心的移动方向，可以有效地减少所需的质心位移。接着使用一种内部交换算法操纵时隙质心。与现有策略（调制数据包的位移）不同，该方法通过在间隔中交换有限数量的 IPD 来改变时隙的质心。实验表明，该方法能有效地抵抗 MFA 攻击，而且在 KLD 测试[22]和 K-S 测试[23]中也无失真现象。

1.3 面向硬件的隐通道和软件网络的流水印

1.3.1 面向硬件的隐通道

ANFW 的研究同时带动了隐通道的研究，两者相互促进。在对于隐通道通信的研究中，Hoda.N 和 Nael.A.G 证明了通用图形处理单元（GPGPU）也能成为隐蔽通信的可行介质[24]。

Chen 等人[25]在 2015 年提出了一个新的微架构级框架 CC-Hunter，它检测共享硬件上可能存在隐蔽定时通道。CC-Hunter 能够成功地在不同的带宽和消息模式下检测不同类型的隐蔽定时信道。Jason 等人[26]在 2014 年提出了一个框架来区分硬件系统中的功能流和时序流，它与门级信息流跟踪（gate level information flow tracking，GLIFT）一起使用，以有效地将时间流与功能流分开，并将其应用在共享总线和缓存嵌入式系统的两个常见资源中。

现代嵌入式计算系统，如医疗设备、飞机和汽车，继续占据本文生活中最重要的一些方面。在这样的系统中，必须严格控制整个设备中信息的移动，以防止隐私泄露，并保证设备安全性。不幸的是，限制信息流通常可能会带来重大挑战，因为信息可能流经难以检测的信道，如定时信道。Venkataramani 等人[27]在 2016 年通过一种微型水平框架检测硬件时间隐通道，从而可以检测到共享硬件上的时间隐通道。Ferraiuolo 等人[28]介绍了定时隔间，这是一种架构抽象，可以在共享的多核处理器上同时运行的软件组之间明确地消除微架构级时序通道。

因此，隐通道在硬件上的研究应用为 ANFW 在硬件上的应用提供了良好的基础，包括框架的设计、载体的选择都启到了一定的推动作用。

1.3.2 面向软件定义的网络流水印

软件定义网络（SDN）通过将控制平面与数据平面分离，为网络运营商提供高度的灵活性和可编程性。用户发起流量时需安装指导流量路由的流规则。该过程需要控制和数据平面之间的通信，使控制器能够监视流量及其来源。

Liu 等人[29]在 2015 年提出了在软件定义网络虚拟环境中的实时定时通道检测。该检测通过框架动态地配置 SDN 以对不同虚拟机（VM）的出站网络流启用/禁用差分分析。文中将框架实现为原型系统，可以在 SDN 环境中动态部署。与现有方法相比，该方法具有更高检测率，可以有效地检测较低延迟和可忽略性能开销的定时信道（TC）。

Park 等人[30]在 2016 年提出了抗免载威胁的流水印方案，一种将水印编码和解码为数据有效载荷的流水印技术。它有助于建立流量规则的所有权，只有流规则的合法所有者可以使用自己的规则发送数据包，网络可以帮助检测已安装的流规则滥用情况。

2　现有网络流水印的攻击方法

在网络安全领域,越来越多的研究人员开始关注匿名通信技术的研究。在某种意义上,匿名通信技术是研究如何抵制各种各样的匿名攻击,其目的是取消系统的匿名性。因此,研究匿名通信技术和通信系统的生存能力将有助于匿名通信系统的发展。

根据攻击的方式,攻击可分为被动和主动攻击。被动攻击者只监视流量和分析相关发送者与接收者之间的沟通信息,而主动攻击者可以修改、增加、删除和数据包延迟。例如,流量分析是被动攻击,拒绝服务攻击是主动攻击。现阶段的主要攻击方式[7,31]有：

a）连续测试攻击。针对网络流水印标识隐形的第一次攻击是 Peng 等人[11]提出的连续测试攻击。设网络中两个节点之间的包延迟序列为 $\delta_1,\delta_2\ldots\delta_N$，一个节点放置在水印机之前，而另一个则被放置在水印机之后。考虑到其中 K 个延时被水印机扩展，为了检测流是否存在水印，令 $\theta = K/N$；同时，将序列概率比测试算法应用于值 θ，结果表明水印被检测出的概率超过 90％以上。

b）多流攻击。Kiyavash 等人[32]提出另一种多流攻击(MFA)方法。攻击者可以发现 Wang 和 Pyun 提出的基于间隔包计数或间隔质心的水印。其中还介绍了如何恢复秘密水印参数。多流攻击基于两个假设：(a)攻击者可以收集少量的水印流，全部包含相同的水印码；(b)正常流可以建模为马尔可夫调制泊松过程（MMPP）。基本上，对于这种攻击，对手可以在时间轴上绘制所有选定流的所有包时间戳，如果流加了水印，则可以在绘图上查看分离好的簇，否则时间戳是均匀分布的。

c）异常通信模式识别。Luo 等人[33]开发了一种水印检测系统，其主要思想是识别网络流中低吞吐量的任何异常序列。该系统针对基于 DSSS 方案扩展的水印进行设计，分为识别流中的低吞吐量周期两个主要步骤；在所选择的时间段内检测异常模式。实验表明攻击者可获得接近 100％的检测率。

Backlit 是针对基于时序的水印进行攻击的例子[34]。与 Luo 等人的工作不同，它通过仅使用非水印通信模式训练的一级分类器来识别异常流量。在训练阶段，算法定义了已知

模式类的边界。分类过程包括确定观察到的流是否是已知类的成员，同时针对 Wang、Houmansadr 和 Pyun 提出的四种水印算法进行了测试。Backlit 框架显示了平均值的异常统计行为，以及基于 RAINBOW[12]和 SWIRL 算法[35]时，发送请求消息与服务器接收响应之间经过的时间方差。Backlit 技术还通过均值平衡算法识别水印插值延迟统计中的异常,这些异常使得 Backlit 可以轻松识别所有测试算法的水印流，精度高达 100%。

d）选择流攻击。Lin 等人[36]在 2012 年的一项工作中说明了一组基于时间的攻击，在他们的模型中称为选择的流动攻击，作者考虑了观测流测量的 IPD 直方图，并对直方图的分区应用余弦相似性度量，以此来揭示水印存在的相似性。攻击实验表明 RAINBOW 和 SWIRL 算法的流水印可以被正确识别，甚至在这种情况下可以获得 100%的精度。

e）均方自相关攻击。Jia 等人[37]最近提出了均方自相关攻击（MSAC）。这种攻击是针对时域和频域的 DSSS 分集方案而设计的。分析证明，当使用相同的 PN 码来扩展水印序列的每个比特时，可以通过识别流中的时序序列相关性。特别是当利用了扩展水印信号的时移周期平方自相关函数时，可以在均方自相关函数中观察到周期性峰值，攻击者能够检测水印的存在并获得 PN 码，其检出率大于 60%。

3　典型 ANFW 的性能对比

前面着重概述了主流 ANFW 算法以及目前存在的一些水印攻击方式，本章在分析了相关文献实验结果的基础上，将从水印算法的隐蔽性、鲁棒性、提取水印时的难易程度（检测率）和实用性等六项指标对其进行对比，如图 4 所示。

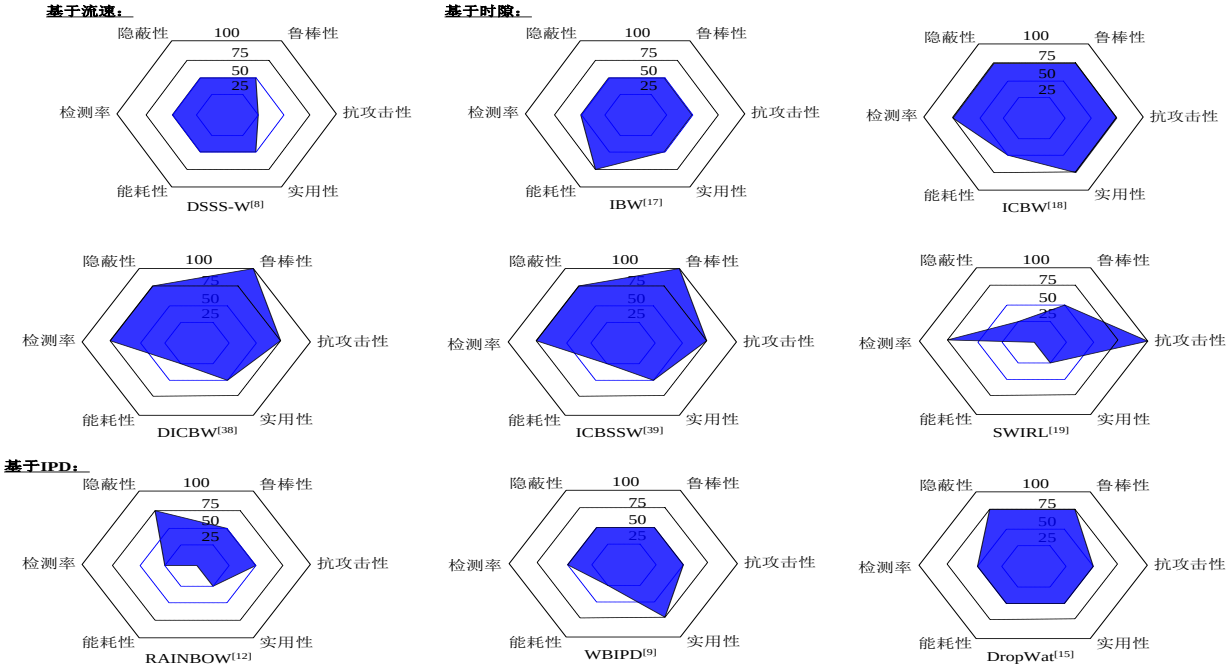

图 4　典型 ANFW 的性能对比

Fig. 4　Performance comparison of typical ANFW

从图 4 中可以直观地了解到这些 ANFW 方法的差别，每一种 ANFW 技术都各有优势与不足。不难看出，基于时隙的流水印方案相比较其他两种方案在性能上更优越，不过带来的缺点就是能耗性非常高，而且其中某些方案性能跨度比较明显，实用性不高。总体看来，基于 IPD 的方案（除 RAINBOW 外）性能都比较平均，实用性也较好。就目前而言，还没有一种完善的水印方案，各有利弊。不过，随着对网络安全的重视，ANFW 技术还会一直处于人们的研究热点当中。

4　结束语

ANFW 技术是数字水印技术与流量分析的一次碰撞。通过在流特征中嵌入特定的可识别的模式（水印），可以使 TA 工具变得更容易和更健壮。本文介绍了网络流水印系统的一般架构，对如今主流的 ANFW 技术进行了分析比较。水印系统的设计中最重要的一个方面就是水印载体的选择，不同的水印载体带来的性能差异较为明显。因此，水印的定位显得尤为重要。Hou 等人[40]提出了一种使用离散余弦变换（DCT）域来定位水印的方法，这是 DCT 在网络水印中的一种新尝试。

由图 4 可以看出主动网络水印技术在性能方面还有待提升，现阶段的主要研究工作依然是完善 ANFW 技术，尤其是在复杂时序信道的出现后[41]，给预防和检测工作提出了更高要求。本文认为今后可从如下方面展开研究工作:

a）提高 ANFW 技术的健壮性。不仅要在网络噪声和抖动等非人为因素的干扰下保持其水印特征，而且在受到网络攻击的情况下，接收方仍能以较大概率正确恢复水印信息，将干扰因素对水印信息的影响控制在固定范围内。在面对不断变化的网络环境和攻击方式时，需要对其不断完善。

b）开发一种 ANFW 的评估模型。可根据每种水印技术的健壮性、隐蔽性、实用性等多个属性量化出一个评估矩阵，再借鉴现有的评估方案、评价矩阵等，给 ANFW 技术一个直观的综合评估，方便安全检测的选择。

c）平衡 ANFW 技术的准确性和隐蔽性。在复杂的网络环境所产生的数据流的差异比较大，对载体的选择/参数的设置提出了更高的要求，在确保高准确性的前提下很难保证水印的隐蔽性。将来研究方向可针对不同类型流研发一种通用机制根据评估模型来选择最适合此流的 ANFW 技术。

d)扩大 ANFW 技术的应用范围和研究领域。目前 ANFW 技术已开始涉及硬件系统、软件应用及云环境下的网络安全检测工作，技术尚未成熟，面对复杂的应用网络稳定性不足，往后需要做大量的实验巩固和完善其在应用网络的有效性和稳定性。随着人工智能时代的到来，伴随着就是人工智能的安全问题，如何在人工智能领域有效地利用 ANFW 技术将是一个重要的研究方向。

e）研究新型流量模式统计分析。设计面向高速网络流的

压缩和统计信息抽取方式，基于压缩感知理论条件，在满足严格的存储空间约束的前提下，设计支持海量数据流并行处理的分析算法，并将传统的流式数据处理拓展到对整个时间轴网络行为的监测，挖掘跨越多个时间段的持久流量模式特征[42]。

参考文献：

[1] Lacovazzi A, Elovici Y. Network flow watermarking: a survey [J]. IEEE Communications Surveys and Tutorials, 2017, 19 (1): 512-530.

[2] Das A K, Pathak P H, Chuah C N, *et al.* Contextual localization through network traffic analysis [C]// Proc of IEEE International Conference on Computer Communications. 2014: 925–933.

[3] Callegari C, Coluccia A, Alconzo A D, *et al.* A methodological overview on anomaly detection [M]// DataTraffic Monitoring and Analysis. Berlin : Springer -Verlag, 2013: 148-183.

[4] Pasqualetti F, Dorfler F, Bullo F. Attack detection and identification in cyber-physical systems [J]. IEEE Trans on Automatic Control, 2013, 58 (11): 2715-2729.

[5] 何高峰, 杨明, 罗军舟, 等. 匿名通信流量在线识别方法 [J]. 软件学报, 2013, 24 (3): 540-556. (He Gaofeng, Yang Ming, Luo Junzhou, *et al.* Tor anonymous communication traffic online identification method [J]. Journal of Software, 2013, 24 (3): 540-556.)

[6] Zander S, Armitage G, Branch P. A survey of covert channels and countermeasures in computer network protocols [J]. IEEE Communications Surveys & Tutorials, 2007, 9 (3): 44-57.

[7] Lu Tianbo, Guo Rui, Zhao Lingling, *et al.* A systematic review of network flow watermarking in anonymity systems [J]. International Journal of Security and Its Applications, 2016, 10 (3): 129-138.

[8] Yu Wei, Fu Xinwen, Graham S, *et al.* DSSS-based flow marking technique for invisible traceback [C]// Proc of IEEE Symposium on Security and Privacy. 2007: 18-32.

[9] Wang Xinyuan, Reeves D S. Robust correlation of encrypted attack traffic through stepping stones by flow watermarking [J]. IEEE Trans on Dependable and Secure Computing, 2011, 8 (3): 434-449.

[10] Wang Xinyuan, Chen Shiping, Aijodia S. Tracking anonymous peer-to-peer VoIP calls on the Internet [C]// Proc of the 12th ACM Conference on Computer and Communications Security. New York: ACM Press, 2005: 81-91.

[11] Peng Pai, Ning Peng, Reeves D S. On the secrecy of timing-based active watermarking trace-back techniques [C]// Proc of IEEE Symposium on Security and Privacy. Washington DC: IEEE Computer Society Press, 2006: 334-349.

[12] Houmanadr A, Kiyavashy N, Borisov N. Rainbow: a robust and invisible non-blind watermark for network flows [C]// Proc of the 16th Network and Distributed System Security Symposium. 2009.

[13] Manjunadha D, Pradeepa S. Enhancing invisibility in network flows by non-blind watermarking technique [J]. International Journal of Research in Computer Applications And Robotics, 2014, 2 (11): 109-114.

[14] 王昌达, 朱慧. 基于包交叉分组间隔质心双盲流水印技术 [J]. 华中科技大学学报: 自然科学版, 2015, 43 (5): 84-88. (Wang Changda, Zhu Hui. Double-blind watermarking technique based on packet cross-closed interval centroid [J]. Journal of Huazhong University of Science and Technology: Natural Science Edition , 2015, 43 (5): 84-88.)

[15] Lacovazzi A, Sarda S, Frassinlli D, *et al.* DropWat: an invisible network flow watermark for data exfltration traceback [C]// Proc of IEEE Transactions on Information Forensics and Security. 2018: 1139-1154.

[16] Lacovazzi A, Sarda S, Elovici Y. INFLOW: inverse network flow watermarking for detecting hidden servers [C]// Proc of IEEE Conference on Computer Communications. 2018: 747-755.

[17] Pyun Y J, Park Y H, Wang Xinyuan, *et al.* Tracing traffic through intermediate hosts that repacketize flows [C]// Proc of the 26th IEEE International Conference on Computer Communications. 2007: 634-642.

[18] Wang Xinyuan, Chen Shiping, Jajodia S S. Network flow watermarking attack on low-latency anonymous communication systems [C]// Proc of IEEE Symposium on Security and Privacy. 2007: 116-130.

[19] Houmansadr A, Borisov N. SWIRL: a scalable watermark to detect correlated network flows [C]// Proc of the 18th Network and Distributed System Security Symposium. 2011.

[20] Lin Zi, Hopper N. New attacks on timing-based network flow watermarks [C]// Proc of USENIX Conference on Security Symposium. 2012: 20-30.

[21] Liu Weiwei, Liu Guangjie, Yang Xia, *et al.* Using insider swapping of time intervals to perform highly invisible network flow watermarking [C]// Proc of Security and Communication Networks. 2018: 1-16.

[22] Archibald R, Ghosal D. A comparative analysis of detection metrics for covert timing channel [J]. IEEE Computer Architecture Letters, 2014, 45: 284-292.

[23] Cabuk S, Brodley C E, and Shields C. IP covert timing channels: design and detection [C]// Proc of the 11th ACM Conference on Computer and Communications Security. 2004: 178-187.

[24] Hoda N, Nael A G. Covert channels on GPGPUs [J]. IEEE Computer Architecture Letters, 2017, 16 (1): 22-25.

[25] Chen Jie, Guru V. CC-Hunter: uncovering covert timing channels on shared processor hardware [C]// Proc of the 47th Annual IEEE/ACM International Symposium on Microarchitecture. 2015: 216-228.

[26] Jason O, Sarah M. Timothy S and Ryan K. Leveraging gate-level properties to identify hardware timing channels [J]. IEEE Trans on Computer-Aided Design of Integrated Circuits and Systems, 2014, 33 (9): 1288-1301.

[27] Venkataramani G, Chen Jie, Doroslovacki M, *et al.* Detecting hardware covert timing channels [J]. IEEE Micro, 2016, 36 (5): 17-27.

[28] Ferraiuolo A, Wang Yao, Zhang Danfeng, *et al.* Full-processor timing channel protection with applications to secure hardware compartments [EB/OL]. [2017-04-25]. https://hdl. handle. net/1813/41218.

[29] Liu A, Chen J, Wrchsler H, *et al.* Real-time timing channel detection in a software-defined networking virtual environment [J]. Intelligent Information Management, 2015, 7 (6): 283-302.

[30] Park Y, Chang S Y, Krishnamurthy L M. Watermarking for detection freeloader misbehavior in software-defined networks [C]// Proc of International Conference on Computing. 2016.

[31] Lu Tianbo, Gao Pan, Du Xiaofeng, *et al.* An analysis of active attacks on anonymity systems [J]. International Journal of Security and Its Applications, 2016, 10 (4): 95-104.

[32] Kiyavash N, Houmansadr A, Borisov N. Multi-flow attacks against network flow watermarking schemes [C]// Proc of the 17th USENIX Security Symposium. Berkeley, CA, USENIX Association. 2008: 307-320.

[33] Luo Xiapu, Zhang Junjie, Perdisci R, *et al.* On the secrecy of spreadspectrum flow watermarks [C]// Proc of the 15th European Symposium on Research in Computer Security. Berlin : Springer, 2010: 232-248.

[34] Luo Xiapu, Zhou Peng, Zhang Junjie, *et al.* Exposing invisible timing-based traffic watermarks with backlit [C]// Proc of the 27th Annual Computer Security Applications Conference. 2012: 197-206.

[35] Houmansadr A, Borisov N. Swirl: a scalable watermark to detect correlated network flows [C]// Proc of Network and Distributed System Security Symposium. 2011: 1-15.

[36] Lin Zi, Hopper N. New attacks on timing-based network flow watermarks [C]// Proc of the 21th USENIX Security Symposium. Berkeley, CA,USENIX Association. 2012: 381-396.

[37] Jia Weijia, Tso P, Ling Zhen, *et al.* Blind detection of spread spectrum flow watermarks [J]. Security and Communication Networks, 2013, 6 (3): 257-274.

[38] Wang Xiaogang, Luo Junzhou, Yang Ming. A double interval centroid-based watermark for network flow traceback [C]// Proc of the 14th International Conference on Computer Supported Cooperative Work in Design. 2010: 146-151.

[39] Luo Junzhou, Wang Xiaogang, Yang Ming. An interval centroid based spread spectrum watermarking scheme for multi-flow traceback [J]. Journal of Network and Computer Applications, 2012, 35 (1): 60-71.

[40] Hou Xueyan, Chen Yonghong, Tian Hui, *et al.* Network watermarking location method based on discrete cosine transform [C]// Proc of the 3rd International Conference on Materials Engineering, Manufacturing Technology and Control. 2016: 1515-1520.

[41] Biswas A, Ghosal D, Nagaraja S, *et al.* A survey of timing channels and countermeasures [J]. Proc of ACM Computing Surveys, 2017, 50 (1): 6: 1-6: 39.

[42] 罗军舟, 杨明, 凌振, 等. 网络空间安全体系与关键技术 [J]. 中国科学: 信息科学, 2016, 46 (8): 939-968. (Luo Junzhou, Yang Ming, Ling Zhen, *et al.* Cyberspace security system and key technology [J]. Chinese Science: Information Science, 2016, 46 (8): 939-968.)