软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2016,27(Suppl.(2)):234−242

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

内容中心网络数据污染的快速检测机制*

汪 漪[1,2], 刘 斌[2]

[1](华为未来网络理论实验室,香港 999077)

[2](清华大学 计算机科学与技术系,北京 100084)

通讯作者: 汪漪, E-mail: wy@ieee.org

摘 要: 内容中心网络通过路由器缓存内容来提高网络的整体性能.为防止被污染的数据在网络中扩散,路由器需要对进入网络的内容进行验证.原始的验证机制需要对内容的数字签名进行非对称密钥解密操作,导致内容验证速度不能满足高速路由器的需求.提出了基于着色的快速内容验证机制,以减少内容验证的计算复杂度,加快内容的检测速度.该机制通过对第 1 次进入网络的正确内容进行着色操作以保证其正确性.被着色的内容再次进入网络时,路由器可以通过着色信息来快速验证内容的正确性,从而提高路由器的检测速度.

关键词: 内容中心网络;数据污染;内容缓存;内容验证;着色

中文引用格式: 汪漪,刘斌.内容中心网络数据污染的快速检测机制.软件学报,2016,27(Suppl.(2)):234−242. http://www.jos.org.cn/1000-9825/16037.htm

英文引用格式: Wang Y, Liu B. Fast content verification for named data networking. Ruan Jian Xue Bao/Journal of Software, 2016,27(Suppl.(2)):234−242 (in Chinese). http://www.jos.org.cn/1000-9825/16037.htm

Fast Content Verification for Named Data Networking

WANG Yi[1,2], LIU Bin[2]

[1](Huawei Future Network Theory Laboratory, Hong Kong 999077, China)

[2](Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: Named Data Networking (NDN) improves the transfer efficiency by caching the contents in routers. To prevent polluted content from being spread in NDN, NDN routers should verify every content that is published in NDN. Since the verification scheme in NDN applies the asymmetric encryption algorithm to sign the content, the verification speed is too slow to satisfy the high speed requirement. This paper proposes a Staining-based verification scheme to improve the verification speed by reducing the computation complexity of the content verification. Staining-based verification scheme stains the content that is sent to the network for the first time; then when the stained content is sent to the network again, the router can utilize the staining information to verify this content. By replacing the asymmetric encryption algorithm with symmetric encryption algorithm, Staining-based scheme can improve the verification speed effectively.

Key words: named data networking; polluted content; cache; content verification; staining content

以 IP 协议为中心的互联网机制已被广泛应用超过了 30 年,这是因为 IP 协议自身的简单性降低了网络互联成本,增强了网络适应性.但从技术发展的角度反思,互联网最初的目标是追求网络的互联以实现硬资源的共享.由于最开始的通信需求发生在两台实体设备间,为确定设备的具体位置,互联网使用 IP 地址来标识不同的设备以支持设备间的数据通信.然而,随着技术的进步和信息化的普及,硬件共享的需求已逐步下降,信息共享成

* 基金项目: 国家自然科学基金(61373143, 61432009); 中国博士后科学基金(2015T80089); 高等学校博士学科点专项科研基金(20130002110084)

Foundation item: National Natural Science Foundation of China (61373143, 61432009); China Postdoctoral Science Foundation (2015T80089); Research Fund for the Doctoral Program of Higher Education of China (20130002110084)

收稿时间: 2016-06-05;采用时间: 2016-10-18

为了主要需求.目前,网络应用的主体已经转为文字、图像和视频信息,内容服务已经成为网络服务的主体.用户关注的不再是内容存储在哪里(where),而是内容本身(what),以及内容检索与传输的速度、质量和安全性.

因此,内容中心网络(named data networking,简称 NDN)[1]被提出并很快成为研究的热点.NDN 直接把内容作为网络处理的基本对象,将数据的存储地址(即当前网络中的节点位置信息)、安全性、可访问性与内容本身分离开来.NDN 的前身是以内容为中心的网络(content centric network,简称 CCN)[2].NDN 与 CCN 的基本涵义一致,都以内容为中心,不再强调含有位置信息的主机,即不再关心内容存储在哪里,而只关心内容的有无和如何快速访问、获取.内容通过名字向网络发布和获取:当主机请求时,它首先向外发送一个数据请求包,该请求包携带被请求内容的名字;一旦能提供被请求内容的节点接收到该请求包,立刻把内容封装并返回给请求节点,完成信息共享操作.从这个过程来看,NDN 和 CCN 并不关心是哪台设备进行了响应,或者设备的位置在哪里,而是关心如何以最小的代价从邻近的节点最迅速地获得该内容,因而有别于传统的 TCP/IP 网络模型.

NDN 网络设备缓存已转发的内容,以减少网络传输带宽、降低网络传输时延、提高网络利用率.由于 NDN 网络设备在响应请求时,直接返回内容,所以在缓存内容时,NDN 网络设备需要验证内容的完整性、正确性,防止被污染的数据在网络中扩散,保证网络安全.数据污染包括如下两种情况.

1) 完整性被破坏的数据,即内容本身与内容携带的签名不一致.例如,恶意用户对内容进行了修改,但由于没有原始内容发布者的私钥,不能再次对数据进行签名,造成数据的不完整.

2) 伪造的数据,即内容本身与内容携带的签名相符,但使用的是错误的私钥进行的签名.例如,恶意用户对内容进行了修改,并使用自己的私钥进行签名,并把公钥、数字签名附带在内容中.

为防止被污染的数据在网络中扩散,NDN 路由器需要对转发和缓存的内容进行验证.基于非对称加密的验证方法(在第 2 节中详细描述)虽然能验证内容的正确性,但由于非对称加密算法的解码过程的速度太慢,不适用于实际的高速 NDN 路由器中.

内容的验证时间主要由两部分构成:对整个内容进行哈希散列计算的时间和对数字签名进行公钥解码的时间,且公钥解码过程消耗了 95.35%的验证时间.在本文中,我们设计、提出了一种基于着色的快速内容验证机制来减少公钥解码的次数,从而提高内容检测的速度.

本文的主要贡献为:

1) 提出了基于消息验证的快速内容验证机制.该机制能够完全避免路由器对内容进行哈希散列计算和非对称加解密操作,但不能防止网络劫持攻击.

2) 提出了基于着色的快速内容验证机制.该机制通过对第 1 次进入网络的正确内容进行着色操作来保证其正确性.被着色的内容再次进入网络时,路由器可以通过着色信息来快速验证内容的正确性.相比 NDN 原始的内容验证机制,基于着色的快速内容验证机制能够提速超过 18 倍.

3) 通过仿真,测试了不同方法的检测速度和传输开销.仿真结果表明,基于着色的内容验证机制能够有效的提高内容的检测速度;在基于着色的内容检测过程中,密文解码过程的时间开销降低为原来的 0.69%.

本文第 1 节对 NDN 内容传输与验证机制进行阐述.在第 2 节中描述基于消息验证的快速内容验证机制.基于着色的快速内容验证机制在第 3 节中描述.第 4 节展示仿真结果.NDN 内容验证的相关研究工作在第 5 节中论述.第 6 节中对本文工作进行总结.

1 NDN 内容传输与验证机制

1.1 NDN数据包类型

与传统的 IP 网络中的单一的数据包类型不同,NDN 网络传输的数据包分为两类:Interest 包和 Data 包.这两种数据包的功能不同,Interest 包是数据请求包;Data 包是数据回复包.用户发送 Interest 包作为对内容(数据)的请求,内容的名字作为最重要的信息包含在 Interest 包内;内容提供者收到 Interest 包后作出回应,先把被请求的内容、内容的名字以及内容提供者自身的信息封装成 Data 包,然后送回给数据请求者.

出于安全和应用程序的需求,Interest 包和 Data 包会额外携带其他信息,其中包括为 Data 包添加数字签名

以认证数据提供者的身份、数字签名的认证机构、数字签名的算法、发布者的基本信息等.总之,除了最基本的要素之外,其他信息可以按需添加到 Interest 包和 Data 包中.

1.2 NDN数据包转发机制

NDN 路由器为了提供更好的服务,存储了 3 种逻辑结构来维护状态信息:内容存储池(content store,简称 CS)、请求状态表(pending interest table,简称 PIT)、转发表(forwarding information base,简称 FIB):

- CS 存储近期被路由器转发的某些常用的数据内容的索引;
- PIT 中存储那些已经被该路由器转发,但还没有收到数据响应的有特殊需求的 Interest 包的状态信息;
- FIB 与 IP 网络中的路由转发表类似,记录 Interest 包可被转发的端口列表;
- NDN 路由器对 Interest 包和 Data 包实行不同的转发机制;
- 当路由器收到 Interest 包时,首先依据 Interest 包中请求的名字作为关键字在 CS 中进行检索,如果存在查询的内容,则直接返回数据给发送请求的用户.如果缓存不命中,则在 PIT 表中进行查找.如果找到对应的表项,说明具有请求相同内容的 Interest 包已经被转发过,但还没有收到 Data 包,所以在对应的 PIT 表表项中添加收到该 Interest 包的端口号,并丢弃该 Interest 包,不再转发;如果没有找到对应的表项,则需要将该名字添加到 PIT 表中,并记录接收到该 Interest 包的端口号,并在 FIB 中进行路由查找.如果 FIB 中没有对应名字的路由信息,则丢弃或返回请求包;
- 当路由器收到 Data 包时,根据 Data 包携带的名字在 PIT 表中进行搜索,获取转发端口列表,然后把 Data 包通过转发端口列表中的端口转发出去,并在 CS 中缓存内容;如果没有搜寻到对应的 PIT 表项,或者 PIT 表项中已经记录同样内容的数据包已经达到,那么丢弃该数据包.

1.3 NDN内容验证方式

NDN 路由器需要对转发和存储的内容进行验证[1],以防止被污染的数据在网络中扩散.图 1 描述了 NDN 路由器进行数据验证的基本过程.

1) 当 NDN 路由器收到 Interest 包对应的数据包时,路由器对 Data 包解析,获取“内容”(D')、“数字签名”($K_{pri}(Hash(D))$)等信息;

2) 路由器对“内容”进行哈希散列计算,获得“内容”的散列值 $Hash(D')$;

3) 路由器使用内容提供者的公钥对数字签名进行解码,获得原始的“内容”的散列值 $K_{pub}(K_{pri}(Hash(D)))$;

4) 路由器将“内容”的散列值 $Hash(D')$与原始的“内容”的散列值 $K_{pub}(K_{pri}(Hash(D)))$进行比较,如果两者相等,则转发并缓存内容 D';否则,丢弃内容 D'.

NDN 路由器中现有的内容验证机制,虽然能够检测到被污染的内容,但验证速度过慢,不能满足高速网络环境的需求.

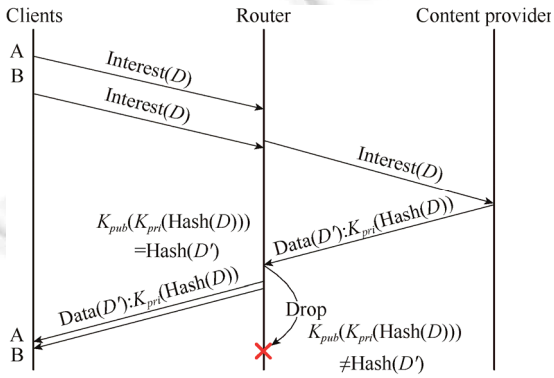

Fig.1 The content verification process in NDN

图 1 NDN 路由器对转发与缓存的内容的验证过程

2 基于消息验证的快速内容验证机制

内容的验证时间主要由两部分构成:对整个内容进行哈希散列计算的时间和对数字签名进行公钥解码的时间.在本节中,我们设计了两种基于消息验证的内容快速验证机制,来避免哈希散列的计算和数字签名的解码操作,从而实现快速的内容检验.

2.1 用户协助的消息验证

如第 1.3 节中所描述的,如果 NDN 路由器计算、解码每个 Data 包中的内容来判断内容是否正确,不仅浪费大量的计算资源,更糟糕的是,现有路由器的计算能力无法满足线速的要求.因此,我们采用对内容提供者(content provider)的身份验证来代替对每个内容的验证.用户协助的消息验证过程如图 2 所示.

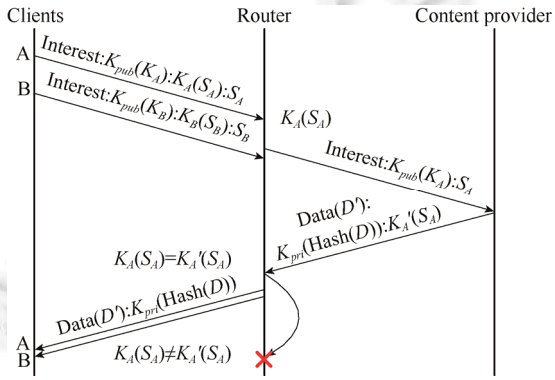

Fig.2 The content verification process with clinets

图 2 用户协助的消息验证过程

1) 用户 A 发送 Interest 包到网络中,同时附带 3 个额外的数据帮助路由器来验证内容提供者的有效性.

a) $K_{pub}(K_A)$,K_A 是临时生成的一个对称密钥.$K_{pub}(K_A)$是使用内容提供者的公钥 K_{pub} 对 K_A 进行加密,它只能通过内容提供者的私钥才能解密得到对称密码 K_A;

b) S_A 是一段随机生成的 64 比特的字符序列,作为原始消息;

c) $K_A(S_A)$是利用 K_A 对 S_A 进行加密后的密文,用于验证内容提供者是否能够正确地解码 $K_{pub}(K_A)$从而获得对称密钥 K_A,即对内容提供者进行身份验证.

2) 当路由器收到用户发送的 Interest 包后,将 $K_A(S_A)$从 Interest 中剥离,并存放在 PIT 中,用于对收到的 Data 包进行验证.剩余的 2 个字段 $K_{pub}(K_A)$和 S_A 则随 Interest 包发送到内容提供者.

3) 内容提供者收到 Interest 包之后,利用私钥 K_{pri} 对 $K_{pub}(K_A)$解析得到 K'_A;然后用 K'_A 对 S_A 加密得到 $K'_A(S_A)$;并将其作为 Data 包的一个字段返回给路由器.

4) 路由器收到 Data 包后,比较存储在 PIT 中的 $K_A(S_A)$和 Data 包中的 $K_A'(S_A)$.如果相等,则表示内容提供者为真实的信息发布者,即表示该内容提供者是可信的.路由器将 Data 包去掉 $K_A'(S_A)$字段后,标注为可信 Data 包,并按照 Interest 包的到达端口返回给请求者;如果不相等,则表示内容提供者不可信,返回的 Data 包中的内容也不可信,直接丢弃 Data 包.

对于相同内容的请求,如图 2 中的用户 A 和用户 B 的行为,虽然用户 A 和用户 B 发送了不同的消息(S_A 和 S_B)和通信密钥(K_A 和 K_B),路由器可以通过请求内容的名字进行识别,并只对上游路由器发送一个 Interest 包,从而减少通信量和内容提供者身份验证的次数.如果返回的 Data 包是可信的,则可以将此 Data 包返回给多个 Interest 包请求者.

由于路由器中缓存的内容都是经过验证的,因此当用户请求的内容已经存储在路由器中时,路由器可以直接返回内容,并保证数据的完整性和正确性.

2.2 路由器自消息验证

在用户协助的消息验证过程中,路由器不需要进行任何的加密、解密操作,仅需要进行消息的比对操作,极大地减少了路由器的计算代价.用户协助的消息验证过程有效地避免了对消息的哈希散列计算及公钥解码过程,使得路由器能够减少这部分计算的资源消耗,并实现线速转发的性能.

然而,在用户协助的消息验证过程中,恶意用户和伪造的内容提供者可以联手发布污染的数据.用户发送一个与伪造的内容提供者串通好的 $K_{pub}(K_A)$,其中 S_A 和 K_A 已被伪造的内容提供者获取;伪造的内容提供者在收到 Interest 包后,向路由器发布任意内容的 Data 包,并将 $K_A(S_A)$作为 Data 包中的验证字段返回.由于路由器存储的验证字段和 Data 包中的验证字段一致,路由器认为内容提供者是可信,从而在转发伪造的内容后,将伪造的内容存储在路由器的缓存中.当其他用户请求相同名字的内容时,路由器会认为存储的内容是真实有效的,将伪造的内容转发给用户.这样,被污染的数据就在 NDN 网络中扩散,失去了内容验证的功能.

因此,我们提出路由器自消息验证机制来防止恶意用户和伪造的数据提供者的串通.图 3 描述了路由器自消息验证过程,不同于用户协助的消息验证机制,路由器担当了消息 S_A、对称密钥 K_A 的生产者.用户按照 NDN 的协议要求,发送 Interest 包到 NDN 网络中,请求名字对应的内容;路由器在收到 Interest 包后,附加上一个 64 比特的消息 S_A 和一个对称密钥 K_A;路由器在收到 Data 包后,解析获得 $K'_A(S_A)$,并将其与 $K_A(S_A)$进行比较.如果相等,则表示内容提供者是可信的;否则,表示内容提供者是不可信的.

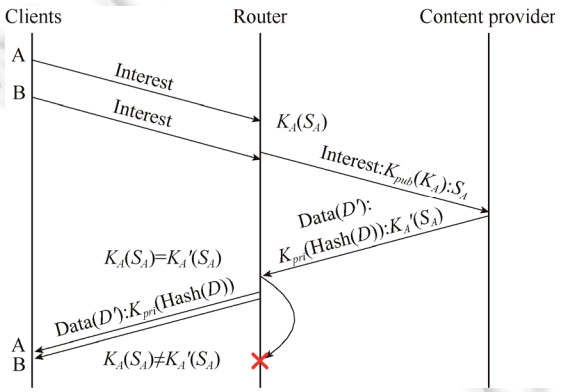

Fig.3 The router cooperation content verification process

图 3 路由器自消息验证过程

与用户协助的消息验证机制相比,路由器自消息验证机制需要路由器生成消息 S_A 和对称密钥 K_A.此外,路由器还需要使用对称密钥 K_A 对 S_A 加密获取消息的密文 $K_A(S_A)$,使用内容提供者的公钥 K_{pub} 加密对称密钥 K_A 获得密文 $K_{pub}(K_A)$.为防止网络重放攻击,路由器需为每个 Interest 请求生成消息 S_i 和对称密钥 K_i.路由器自消息验证机制对每个到达的 Data 包,虽然节约了一次对内容的哈希散列计算,但至少需进行一次在线非对称加密操作,这就使得路由器自消息验证机制需要使用大量的计算资源来保证线速的内容验证.

消息 S_i 和对称密钥 K_i 与 Interest 包和 Data 包无关,路由器可以预先或离线生成 S_i、K_i、$K_i(S_i)$以及 $K_{pub}(K_A)$,从而减少路由器在内容验证环节的计算复杂度,提高路由器内容验证的速度.但采用离线方式生成验证信息时,会使得验证信息与数据本身没有任何的关联,造成路由器不能防止网络劫持和重放攻击.

3 基于着色的快速内容验证机制

虽然路由器自消息验证机制能够防止被污染数据在 NDN 网络中传播,但会降低 NDN 网络的整体性能.在基于消息验证的快速内容验证机制下,用户仅能够获取由内容提供者发送的内容或已经缓存在路由器中的内

容,不能获取网络中其他用户缓存的内容,未能发挥 NDN 网络的缓存共享的特性,从而造成 NDN 网络传输效率的降低.另一方面,由于基于消息验证的内容验证机制不会检测消息本身,不能防止内容劫持和重放攻击.

因此,我们提出基于染色的快速内容验证机制来保持 NDN 网络中内容共享的特性,并减少同一内容在同一自治域(autonomous system,简称 AS)内的路由器中的内容检测次数,从而加快内容检测的速度.基于染色的快速内容验证机制的具体过程如图 4 所示.

1) 用户 A 的 Interest 包转发到内容提供者后,内容提供者发送 Data 包给路由器 1,其中包括内容 D'和对内容的散列值进行私钥加密后的密文 $K_{pri}(Hash(D))$.

2) 路由器 1 对收到的内容 D'进行哈希计算得到散列值 $Hash(D')$,并对内容提供者的数字签名进行解码获得原始内容的散列值 $K_{pub}(K_{pri}(Hash(D)))$.

3) 路由器 1 比较 $Hash(D')$和 $K_{pub}(K_{pri}(Hash(D)))$,如果不相等则丢弃;如果相等则转发内容到用户 A,并在 Data 包中附加路由器 1 对内容的散列值的对称加密信息 $K_R(Hash(D))$.与此同时,路由器 1 将 $K_R(Hash(D))$附加在 Interest 包中发送给内容提供者.内容提供者在收到携带有 $K_R(Hash(D))$的 Interest 包后,存储 $K_R(Hash(D))$到内容 D 的附件数据中,以便下次直接发送 $K_R(Hash(D))$给相同自治域内的路由器.其中,我们称 $K_R(Hash(D))$为对内容 D 的染色标签.

4) 当用户 B 发送对内容 D 请求的 Interest 包到路由器 2 时,路由器 2 发现用户 A 有内容 D 的缓存,所以转发 Interest 包到用户 A.

5) 用户 A 收到 Interest 包后,发送内容 D 及先前收到的对内容 D 的染色标签 $K_R(Hash(D))$给路由器 2.

6) 路由器 2 收到内容 D'和 $K_R(Hash(D))$后,使用对称密钥 K_R 对 $K_R(Hash(D))$进行解码获得 $K_R(K_R(Hash(D)))$.

7) 如果 $Hash(D')$与 $K_R(K_R(Hash(D)))$相等,路由器 2 将内容 D'发送给用户 B;否则,丢弃内容 D',并发送 Interest 包给上游路由器或其他用户请求内容.

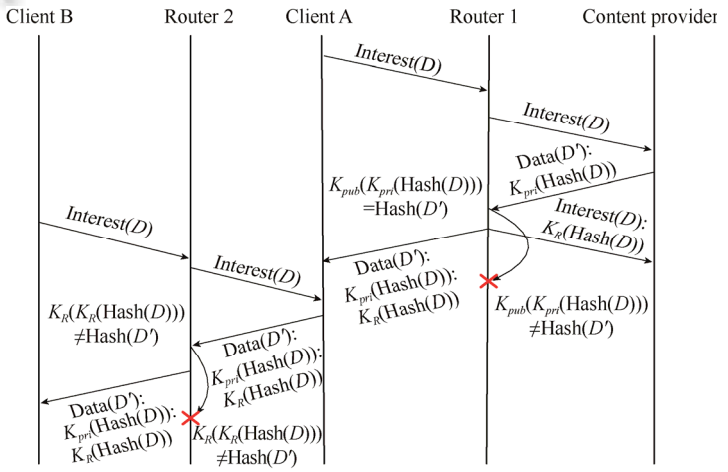

Fig.4 The content verification with staning mechanism

图 4 基于着色的快速内容检测过程

基于着色的快速内容验证机制适用于同一自治域内的路由器相互协助进行内容验证.相比基于消息验证的内容检测,具有以下的特点:

- 一个内容在一个自治域内仅需要一次非对称解密操作.相比原先的 N 次非对称解密操作(N 为内容发送的次数),可以减少 N–1 次非对称解密操作的计算开销.
- 同一自治域内的不同路由器协同工作以完成内容的验证.所有路由器使用相同的对称密钥 K_R 对经过验证的内容进行染色,以便在同一自治域内进行传播和快速验证.

- 缓存在用户中的内容可以被其他用户访问,充分利用了 NDN 网络的内容共享的特性,提高了 NDN 网络的传输效率.

网络攻击者可试图通过破解对称密钥 K_R 来污染内容.为防止这种攻击,基于着色的快速内容验证机制,不定时的修改 K_R,从而避免这种针对 K_R 的攻击.

4 仿真结果

本节中,我们通过系统仿真来测试上述 4 种验证机制的性能,主要从验证速度和传输开销两个方面来评价其优劣.4 种机制分别是:NDN 原始内容验证机制(NDN-verify)、基于用户协助消息验证的快速内容验证机制(user-verify)、基于路由器自消息验证的快速内容验证机制(router-verify)以及基于着色的快速内容验证机制(staining-verify).

在仿真系统中,各种基本的加密算法的性能见表 1 和表 2.各种算法使用 C++构造代码,用 Visual C++ 2005 SP1 编译,在一台装有 Windows Vista 32-bit 操作系统的普通服务器上进行测试.该服务器的 CPU 型号为 Intel Core 2 1.83GHz.在运行程序时,关闭 OpenMP 功能,使得每个程序都在单线程模式下运行.

Table 1 The performance of hash functions and symmetric encryption algorithms

表 1 哈希散列函数和对称加密算法的性能

Algorithm	MB per second	Cycles per byte
CRC32	253	6.9
MD5	258	6.8
SHA-256	111	15.8
DES/CTR	32	54.7
DES-XEX3/CTR	29	60.6
DES-EDE3/CTR	13	134.5

Table 2 The performance of asymmetrical encryption algorithms

表 2 非对称加密算法的性能

Algorithm	Operations per second	Cycles per operation
RSA1024 Encryption	12 500	140 000
RSA1024 Decryption	684.93	2 680 000
RSA2048 Encryption	6 250	290 000
RSA2048 Decryption	164.47	11 120 000

表 1 所示哈希散列算法中,CRC32 和 MD5 具有较好的性能,每秒约可以处理 250MB 的数据.另一方面,CRC32 只能生成一个 32 位的散列值,MD5 能生成一个 128 位的散列值.DES 算法统一使用 CTR 模式,DES、DES-XEX3 和 DES-EDE3 分别可以实现 32MB、29MB 和 13MB 的加密速度.综合考虑,我们选用 MD5 作为哈希散列计算函数、DES-XEX3 来实现对称加密.

如表 2 所示,非对称加密算法 RSA 比对称加密算法 DES 要慢 2~3 个数量级.RSA 的加密速度会比其解密速度慢 1~2 个数量级.在表 2 中,RSA1024 的加密速度为每秒 12500 次,而 RSA1024 的解密操作为每秒 684.93 次;RSA2048 的加密速度是 RSA1024 的 50%,但解密速度却是 RSA1024 的 24%.这就说明,随着 RSA 密钥长度的增加,RSA 的加密与解密的速度的差距会越来越大.

4.1 内容验证速度

内容验证的时间由两部分组成:计算内容的哈希散列的时间和加解密的时间.计算内容的哈希散列的时间会随着内容大小的增加而增长;加解密的时间会因为采用的加解密方法不同而变化.因此,在本小节中,我们给出了不同情况下,各种算法的验证速度.其中,哈希散列计算采用 MD5 算法,非对称加密采用 RSA1024 算法,对称加密采用 DES-XEX3/CTR 算法.

表 3 列出了不同方法在验证大小为 1 000 Byte 的内容时花费的时间.NDN-Verify 方法需要 0.003 9ms 来计算哈希散列,0.08ms 来验证数字签名,共花费 0.083 9ms.User-Verify 方法将所有的计算开销都留给了用户和内容

提供者,路由器不需要进行任何的计算,因此其花费的时间为 0ms.Router-Verify online 方法虽然不需要对内容进行哈希散列计算,但需要在线实时的计算 $K_A(S_A)$和 $K_{pub}(K_A)$,因此需花费 1.46ms 来验证一个内容.Router-Verify offline 方法离线生成 $K_A(S_A)$和 $K_{pub}(K_A)$,减少了 Router-Verify online 方法中计算 $K_A(S_A)$和 $K_{pub}(K_A)$的时间开销,因此路由器不再需要进行任何计算,花费的时间为 0ms.Staining-Verify 方法对第 1 次进入网络的内容需要进行散列计算和解码数字签名,因此需要花费与 NDN-Verify 相同的时间 0.0839ms.但当同一内容再次进入相同的 NDN 网络时,由于已经被染色,因此可以通过验证着色信息来验证内容,内容的验证时间为 0.00445ms.总的来说,在再次验证大小为 1000 字节的内容时,Staining-Verify 方法的效率是 NDN-Verify 方法的 18.85 倍.

Staining-Verify 方法在再次验证内容时,主要时间开销由原来的非对称解密操作的时间变为对内容的哈希散列计算的时间.哈希散列的计算与加解密计算分别占用 87.64%和 12.36%的时间.因此,在实际的路由器实现中,需使用硬件加速哈希散列的计算时间,从而提高路由器的内容验证性能.

Table 3 The speeds for a content with 1 000 Bytes

表 3 验证一个大小为 1 000 字节的内容的时间

Scheme	Hash time (ms)	Encryption and decryption time (ms)	Total time (ms)
NDN-Verify	0.003 9	0.08	0.083 9
User-Verify	0	0	0
Router-Verify (online)	0	1.46	1.46
Staining-Verify (First time)	0.003 9	0.08	0.083 9
Staining-Verify (Not first time)	0.003 9	0.000 55	0.004 45

4.2 传输开销

不同的验证方法,需要在原始的内容中附加不同的数据,因此需要不同的传输开销.所有方法的网络开销见表 4,其中 M 表示用户到内容提供者或内容缓存之间的网络跳数.NDN-Verify 方法仅需要 64Byte 的数字签名(内容散列值经过私钥加密后的密文),因此经过 M 跳后,总共传输了 64M 字节.User-Verify 方法需要附加消息 S_A、及对称密钥 K_A 在公钥加密后的密文 $K_{pub}(K_A)$在 Interest 包中,因此总共传输了 144M+8 字节,是 NDN-Verify 方法的 2.25 倍.Router-Verify 方法仅在最后一个路由器向内容提供者发送的 Interest 消息中附加验证消息,因此仅需要 88+64M 字节.Staining-Verify 方法对内容的第 1 次验证后,会额外附加着色信息(8 字节)到 Data 包中,因此总共传输约 72M 字节,相比 NDN-Verify 方法,仅有 12.5%的额外开销.

Table 4 The overheads of different mechanisms

表 4 不同方法的传输开销

Scheme	Message (Byte)	Ciphertext (Byte)	Total (Byte)
NDN-Verify	0	64M	64M
User-Verify	8M	136M+8	144M+8
Router-Verify	8	80+64M	88+64M
Staining-Verify (First time)	0	64+72×(M–1)	72M–8
Staining-Verify (Not first time)	0	72M	72M

5 相关工作

目前,学术界和工业界对 NDN 的研究,主要集中在快速名字查找[3−7]、高效的内容缓存策略[8,9]、快速转发机制[10]、基于 NDN 的应用[11,12]等.前期的这些研究论证了 NDN 技术的可行性和实用性,为推动 NDN 的发展做出了贡献.

NDN 的安全机制,作为 NDN 的重要组成部分,也获得了广泛的关注.现阶段,对 NDN 的安全机制的研究,主要集中在用户的隐私保护[13,14]、检测缓存失效攻击[15,16].用户的隐私保护的提出是由 NDN 网络会缓存内容引起的,即一个用户请求的数据会暂存在路由器上,恶意用户可以通过请求数据的响应时间来判断哪些内容是其他用户感兴趣的,从而获取其他用户的隐私信息.文献[13,14]通过增加响应时间使得根据内容响应时间来判断用户请求的隐私探测机制失效,从而有效的保护用户的隐私.文献[15,16]提出了多种机制来检测用户发送的请求的次数和流行程度,从而判断用户的行为.当判定用户行为为恶意行为时,不再转发该用户的 Interest 包,从而

避免缓存失效.根据我们的调研结果,本文首次提出基于着色的方法来加速内容的验证过程,从而减少计算代价,提高 NDN 路由器的性能.

6 结束语

本文描述了基于消息验证的快速内容验证机制和基于着色的快速内容验证机制.基于消息验证的快速内容验证机制能够完全避免路由器对内容进行哈希散列计算和非对称加解密操作,但不能避免网络劫持攻击且不能有效的利用 NDN 的缓存策略.基于着色的快速内容验证机制通过对第 1 次进入网络的正确内容进行着色操作来保证其正确性.被着色的内容再次进入网络时,路由器可以通过着色信息来快速验证内容的正确性.相比原始的 NDN 内容验证机制,基于着色的快速内容验证机制能够提速约 20 倍.仿真结果表明,基于着色的内容验证机制能够有效的提高内容的检测速度.

在基于着色的内容检测过程中,密文解码过程耗时降低为原来的 12.36%.使得内容的哈希散列计算成为性能的瓶颈,占用了 87.46%的计算时间.因此,为进一步提高内容验证的速度,路由器需使用专用硬件来实现对内容的哈希散列计算.

References:

[1] Named data networking (NDN) Project. 2010. http://www.named-data.net/techreport/TR001ndn-proj.pdf
[2] Jacobson V, Smetters DK, Thornton JD, Plass MF. Networking named content. In: Proc. of the 5th Int'l Conf. on Emerging Networking Experiments and Technologies (CoNext). 2009.
[3] Wang Y, He KQ, Dai HC, Meng W, Jiang JC, Liu B, Chen Y. Scalable name lookup in NDN using effective name component encoding. In: Proc. of the 32nd Int'l Conf. on Distributed Computing Systems (ICDCS). 2012.
[4] Wang Y, Zu Y, Zhang T, Peng KY, Dong QF, Liu B, Meng W, Dai HC, Tian X, Xu ZH, Wu H, Yang D. Wire speed name lookup: A GPU-based approach. In: Proc. of the 10th USENIX Symp. on Networked Systems Design and Implementation (NSDI). 2013.
[5] Wang Y, Pan T, Mi Z, Dai HC, Guo XY, Zhang T, Liu B, Dong QF. NameFilter: Achieving fast name lookup with lowmemory cost via applying two-stage Bloom filters. In: Proc. of the 32nd Annual IEEE Int'l Conf. on Computer Communications (Infocom). 2013.
[6] Wang Y, Dai HC, Jiang JC, He KQ, Meng W, Liu B. Parallel name lookup for named data networking. In: Proc. of the IEEE Global Telecommunications Conf. (Globecom). 2011.
[7] Wang Y, Xu B, Tai D, Lu JY, Zhang T, Dai HC, Zhang BC, Liu B. Fast name lookup for named data networking. In: Proc. of the 22nd Int'l Symp. of Quality of Service (IWQoS). IEEE, 2014. 198–207.
[8] Wu H, Li J, Wang Y, Liu B. EMC: The effective multi-path caching scheme for named data networking. In: Proc. of the 22nd Int'l Conf. on Computer Communication and Networks (ICCCN). IEEE, 2013. 1–7.
[9] Rosensweig EJ, Menasche DS, Kurose J. On the steady-state of cache networks. In: Proc. of the IEEE INFOCOM. 2013. 863–871.
[10] Cheng Y, Afanasyev A, Moiseenko I, Wang L, Zhang BC, Zhang LX. A case for stateful forwarding plane. Computer Communications, 2013,36(7):779–791.
[11] Jacobson V, Smetters DK, Briggs NH, Plass MF. VoCCN: Voice-Over content-centric networks. In: Proc. of the 2009 Workshop on Re-Architecting the Internet (ReArch). 2009.
[12] Zhu ZK, Wang S, Yang X, Jacobson V, Zhang LX. ACT: Audio conf. tool over named data networking. In: Proc. of the ACM ICN Workshop. 2011.
[13] Acs G, Conti M, Gasti P, Ghali C, Tsudik G. Cache privacy in named-data networking. In: Proc. of the ICDCS. 2013.
[14] DiBenedetto S, Gasti P, Tsudik G, Uzun E. ANDaNA: Anonymous named data networking application. In: Proc. of the NDSS. 2012.
[15] Xie MJ, Widjaja I, Wang H. Enhancing cache robustness for content centric networks. In: Proc. of the INFOCOM. 2012.
[16] Conti M, Gasti P, Teoli M. A lightweight mechanism for detection of cache pollution attacks in named data networking. Computer Networks, 2013,57(16):3178–3191.

汪漪(1983－),男, 浙江杭州人,博士,CCF 会员,主要研究领域为内容中心网络,软件定义网络,高性能网络架构与设备.

刘斌(1964－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为内容中心网络,软件定义网络,高性能网络架构与设备.