第7章 CHAPTER

可追查性检查思想

笔者在对于邮件协议、smtp 协议漏洞、垃圾邮件产生及泛滥原因等进行分析后，考察了大量的垃圾邮件的发送路径、邮件头信息，认为基于黑名单（包括动态黑名单 RBL）及内容分析等反垃圾邮件方案偏离了垃圾邮件产生的根本原因，无法根除垃圾邮件。在考虑到 smtp 服务器已经遍及全球的事实情况下，于 2001 年初提出了如下方案，命名为“可追查性检查”方案，但直到 2003 年初，才有人使用该方案产生了产品或雏形产品。

可追查检查方案是一个保证收到的邮件都是可以追查来源的方案。对于无法追查来源的邮件，用户无法回复（这些垃圾邮件也无需用户回复）、公安监管无法查找（使监管人员被大量无法追查的邮件淹没）、加大了反垃圾邮件其他技术的负荷（内容智能分析很耗系统资源）。

当然，存在未做伪装来源的垃圾邮件，但这种垃圾邮件经常是广告等发件人愿意负责的邮件，这样的垃圾邮件实际上定义不清，有些人认为是垃圾，有些人可能认为不是垃圾。这种垃圾邮件可以靠已有防垃圾技术（如黑名单）、投诉机制、法律法规等方法解决。

不可追查的邮件是伪造了邮件作者、发件人、邮件转发路径等信息的邮件，使接收人无法根据邮件找到发件人。原先的不可追查邮件曾使用不存在的域名，这样，有些 MTA 在接收邮件时，增加了域名解析，如果发现发件人域名不存在，则拒收邮件，这曾很大程度上阻止了垃圾邮件。

而如今的垃圾邮件（和病毒邮件）都使用存在的域名，病毒邮件经常还是在被感染者的通信簿中寻找发件人和收件人邮件地址，这样使垃圾邮件和病毒邮件更具迷惑性（因为收件人会认为邮件是从他所熟悉的发件人发来的）。

这样的垃圾邮件还导致一个恶劣后果，就是被冒用的发件人邮件地址可能要承担垃圾邮件导致的后果，例如可能被受害的收件人抱怨、被列入黑名单，甚至面临法律问题。

7.1 基本理念

- 把关注重点放在判断含有虚假的信息源、发件人、路由等信息的垃圾邮件上；
- 重新定义垃圾邮件：做了任何伪装的、不愿承担邮件后果的邮件是垃圾邮件；反之：负责任的邮件服务器上的负责任的合法用户发出的邮件为合法邮件。
- 有效性和易实施性的统一；
- 不企图追查所有邮件的来源，拒绝不可追查的邮件，只接收可追查的邮件，如需追查来源也只追查可追查的邮件。

凡是不符合诚信规定的邮件和服务器，我们都认为它是垃圾邮件予以拒收。

我们不认为所有的广告邮件都是垃圾邮件，如果广告邮件出自负责任的服务器和用户，可以知道该服务器的固定 IP 地址和托管地点，如果不厌烦则允许接收，如果厌烦了，只需要增加黑名单即可。

7.2 理想

一个理想的反垃圾邮件系统应该具备以下特点。

- 技术简单：最好与现有技术兼容。
- 性能优异而稳定：达到摒除大部分垃圾邮件的效果。
- 易于实施：无需或尽量少其他邮件服务器配合，哪里实施、哪里有效。
- 管理容易：不应增加管理员太多工作，不应要求其他邮件域管理员做某些工作。
- 可逐步实施、自愿实施，不单独靠行政指令。

7.3 希望与现实

用户通常希望一封邮件所声称的发送者(Mail From：)是真正的发送者，而不是伪造的，那么应该满足以下条件

- 发送邮件的计算机的 IP 应该是“Mail From：”的邮件域中允许以该域名义发送邮件的计算机。
- “Mail From：”应该和邮件内容中的“From：”相同，下面统称发件人。

- 发件人应该是发件域的合法用户。
- “Rcpt To：”应该可以和“To：”不同，如多邮件接收者。

遗憾的是，互联网上目前的服务和协议根本没有合法发件人和合法发件计算机的数据，另外牵涉到邮件用户隐私、安全保护等因素，连域名下的发送邮件的 SMTP 服务器也经常被隐藏；合法邮件用户地址按照反垃圾邮件的原则，应尽量隐藏以免被垃圾软件利用；甚至 SMTP 协议允许空发件人，以发送错误反馈邮件。

7.4 第一步实现

在考察互联网邮件服务器实际运行方式的前提下，有如下假设和实验。

- 大量的接收邮件服务器（域名解析的 MX 项）本身就是发送邮件服务器（尤其针对中小邮件服务器）。
- DNS 解析给出的 IP 地址一般是受域拥有者控制的，因此允许发送邮件。许多 DNS 对域名本身也做了 A 记录解析（域名本身是台机器的名字）。
- 大部分分离的接收邮件服务器和发送邮件服务器是 IP 地址相邻相近的，可以通过范围扩展，在 MX 所指定的接收邮件服务器附近的 IP 地址都认为可能是相关的、合法的发送邮件服务器。
- 其他可能允许发送邮件的计算机，例如 www/mailist/ftp 服务器，可能与 MX 项定义的计算机的 IP 地址相邻相近。
- 部分邮件服务器具有 DNS 反解功能，反解得到的域名和用户声称的相同（尤其针对大型邮件服务器或跨国公司）。但应注意：DNS 反解是可以被垃圾邮件发送者伪造的，因此使用 DNS 反解需要慎重。
- 为支持错误回答邮件的空“Mail From：”，可以用 HELO/EHLO 参数代替“Mail From：”的邮件域部分，基本可以达到相同的效果。
- 从 Mail From：得到的一般是域名，从 HELO 得到的经常是计算机名，但通过 DNS 解析，可以区分和提取。

基于上述假设，基本可以正确地判断一封邮件是否来自其声称的邮件域的合法计算机。

第一步实现如下效果。

- 有效性高：可以很容易摒除 95％的垃圾邮件，屏蔽了绝大部分的邮件病毒。
- 兼容性强：可以和其他防垃圾方案配合使用，且大幅度提高其他方案的有效性。
- 合作性强：可以和其他没有防垃圾措施的邮件系统配合使用，不要求其他邮件服务器（发件方）做任何工作。

- 高性能、易实施。

但还是是存在漏判和误判的问题。

漏判和误判

漏判的原因是基于相邻假设而扩大匹配范围。

误判的原因是实际情况不完全满足上述假设。

误判补偿包括降低严格度(将范围扩大到一个 B 类地址仍能保证 95%以上的垃圾邮件被排除)和网站补偿。

7.5 内部可追查检查

目前,关闭了开放式转发功能、对外发邮件需要用户认证的邮件系统存在两个问题。

- 用于认证的用户名/口令可以和邮件本身的 From、Mail From 不同。
- 向本服务器用户发送的邮件无需认证。

第二个问题实际上是发件 MUA 直接向收件 MTA 发送邮件,和发件 MTA 向收件 MTA 发送邮件的过程完全相同,但外部可追查检查不适合解决这种情况,因此要求同一邮件服务器的合法用户向另一用户发送邮件时,也必须使用认证,否则就认为是垃圾邮件。其他域的用户直接向本域收件 MTA 发邮件时,应该使用外部可追查解决。

外部可追查检查用来保证发件 IP 地址属于邮件所声称的邮件域,如果发件人的邮件服务器不是开放式转发的(即外发邮件需要认证),则发件人必然是该邮件域的合法用户,但无法保证他以其在该邮件域的真实邮件地址发送邮件,虽然在发送认证时他需要提供该邮件服务器的合法用户名和口令,但他完全可以在邮件的“Mail From:”和“From:”中提供假的信息,导致收件人只能知道该邮件是发件域的合法用户发出的,而无法知道是谁。这就是上面提到的第一个问题。

内部可追查检查要求无论向哪里发邮件都必须使用发件认证,且邮件的“Mail From:”与发送认证用户必须匹配,这样保证发出的邮件可以追查到发件人。该方案虽然简单,但可以做到类似于邮件签名的一些特点,从某种意义上考虑,可以相信该邮件不是别人伪造的。

实施内部可追查检查的前提是关闭开放式转发,外发邮件需要认证。

7.6 全面可追查检查

上述可追查方案的第一步实现，是基于无需无关者修改任何程序、配置，甚至无需知道的前提下，实施了外部可追查检查的用户可以保证拒绝大部分不可追查的垃圾邮件，实施了内部可追查检查的用户可以保证防止内部合法用户伪装别人外发垃圾邮件。

基于客观的情况（互联网上无法得到需要的认证信息），因此上述第一步实现必然存在漏判和误判。在上述误判补偿中，使用了网站补偿，将不符合假设的域名用网站来补偿。那么，如果将网站补偿的方案推而广之，就可以产生全面可追查检查方案，完全防止以伪装方式发出的垃圾邮件和病毒邮件。

也就是说，需要在互联网上提供一套发件人认证的机制，来确认发件人是从其声称的邮件域所许可的 IP 地址的计算机发出的邮件，进而确认发件人确实是该邮件域的合法用户。当发件 MTA 通过 SMTP 协议向收件 MTA 发送邮件时，收件 MTA 可以通过这套发件人认证机制，确认发件 MTA 拥有以其声称的邮箱发送邮件的权力，这时才予以接收。

这套机制就实现上，可能使用集中式和分布式两种方案。

- 集中式：类似域名注册机构的管理，由特定的机构维护合法发件 IP 地址或发件人数据，这样的方案是中心化管理，有利于统一管理，但实施上有些难度，尤其是全球化实施。另外的问题可能是管理机构维护有一定的难度，但这部分难度可以靠将全集中分散为树枝型来部分解决。
- 分布式：在与域名、邮件服务器等服务密切相关的服务程序上提供改进方法，可能是修改服务程序，也可能是利用现有协议和程序增加配置，来完成可追查性检查的认证。最可能利用的就是 LDAP 和 DNS，LDAP 可以提供更灵活的应用和更完全的支持，但不是所有域都有 LDAP 服务器，甚至可以说很少。DNS 服务器是每个域都必须有的功能，最容易利用。当可追查检查的方案被逐渐推广时，不提供可追查检查认证数据的域就有可能被误拒绝，这就可能促使该域的管理员作出相应的调整和配置。

具体使用哪种方法，应考虑到实施难度及覆盖范围、用户接受难度及标准化等方面的因素。应该让域名拥有者管理自己的事，而不必有交叉工作。要考虑域名注册商、DNS 服务提供商、主机托管商、主管机构的相关工作。

7.6.1 要考虑的认证数据

1. 主要机制

- 某个邮件域允许发送邮件的 IP 地址、IP 子网或 IP 范围列表：发送邮件的计算机的 IP 地址应在该邮件的 Mail From 用户的邮件域允许发送邮件的 IP 地址列表中。
- 某个邮件域可能为其他邮件域提供邮件转发服务：发送邮件的计算机的 IP 地址所在的域应明示提供给该邮件的“Mail From：”用户的邮件域提供邮件转发服务，重定向机制。
- 应考虑父域可以为子域提供可追查认证的机制，即当可追查检查客户端无法在其域名得到认证服务时，可以向其父域发起认证请求：递归/重定向机制。
- 支持动态 IP 地址：应提供动态 IP 地址的合法邮件发送机给该 Mail From 用户与该域的可追查认证服务动态认证的机制，保证该邮件可以被可追查检查认证服务客户端（收件 SMTP 服务器）接收，例如可以利用该邮件的 Message-ID 等信息，或者使用动态 DNS 方式，在可追查检查认证服务中记录主机名，而给可追查检查客户端提供当时的动态 IP 地址。
- 同时支持 IPv4 和 IPv6。

2. 次要机制

- 为方便管理，可能需要提供“排除机制”，即上述定义的非逻辑。
- 为了简单，可以提供 DNS 服务所解析的所有 MX 记录（在 7.4 节实现中已实现）、A 记录等作为可发送邮件服务器的 IP 地址。
- 为了简单，可以提供 LDAP 服务所提供的所有合法用户，并以某种认证方式保证 LDAP 认证用户当时使用的 IP 地址的合法性。
- 为了简单，可以要求 DNS 配置反向解析，这样可以初步实现认证机制。
- 递归授权。

7.6.2 可得到的效果

采用全面可追查检查的方案后，无需范围扩展，因此不会发生漏判；由于所有法定允许发邮件的计算机应该在可追查检查认证服务中定义，因此不会发生误判。

当然，该方案的实施肯定是渐进的，因此在可追查检查客户端可以结合 7.4 节的第一步实现和全面可追查方案：当邮件域没有提供可追查检查认证服务时，使用原来的第一

步方案。

结合 RBL 等技术和相关法律法规，可以基本做到“彻底”解决垃圾邮件问题。

要注意的是：全面可追查方案和前面的可追查第一步实施有相互冲突之处，因此设计时应在没有全面可追查方案认证数据的情况下，才使用可追查第一步实施；如果有了全面可追查方案认证数据，就不再做第一步实施的检查。

7.7 可追查检查实施方案

可追查检查可以采用分布式方案，类似上述 DNS 增强方案，同时，也可以采用集中式方案，类似域名注册服务。但这两种方案各有优缺点。集中式方案在管理上存在一定的难度。

分布式方案则存在伪装的可能性。例如垃圾邮件发件人使用合法注册的域名发送垃圾邮件，这不会被可追查检查认为是非法，但可能被某种黑名单阻止。但假如垃圾邮件发件人注册许多域名，如 1000 个，每个域名每天发送的垃圾邮件数量很少，可能不会被黑名单列入。

这时，需要集中和分布相补充的方案。集中方案：负责对分布许可的可靠性认可；分布方案：负责本域信息发布。

在可追查中心登记的域名，将对此域名发送的邮件负责，可追查检查系统完全相信该域名的分布可追查服务：可追查服务认为是垃圾邮件的就是垃圾邮件，认为不是垃圾邮件的就不是垃圾邮件。

对于未在可追查中心登记的域名，可追查检查系统部分相信该域名的分布可追查服务：可追查检查认为是垃圾邮件的就是垃圾邮件，认为不是垃圾邮件的可能还是垃圾邮件（有可能是垃圾邮件发送者注册的域名）。

7.8 该方案应用及国内外类似方案

可追查性检查方案已经在国内近十家的反垃圾邮件产品中应用或即将应用，其中已取得公安部销售许可证的厂家有 4 家。

类似的方案，在国内一直无人提出，亦无人重视，甚至笔者与一些同业者交流、探讨后，依然有许多人不理解其意义所在。

类似技术中，名声最大的当属微软公司的 SenderID，2004 年 7 月 Bill Gates 来北京，

宣称微软有最先进的反垃圾邮件技术，同时微软向 IETF 提交了 RFC 建议标准的申请，但被 IETF 拒绝。

SenderID 使用 XML 格式定义，并兼容 SPF。SPF 是新加坡 Pobox 公司于 2003 年 12 月 10 日公布的类似方案，目前已经有一定规模的使用。

国外最早的类似技术是德国人于 2002 年 12 月提出 RMX(反向 MX)方案，目的就是要指出一个邮件域的合法发件服务器。

7.9 可追查性检查与反垃圾邮件立法

1. 立法实施的技术问题

在前面我们已经分析过，反垃圾邮件立法存在实施难度。对于种种进行过地址伪装，因而导致不可追查的垃圾邮件，不仅占用了大量网络带宽，而且使用内容监控提交管理机关时，会导致大量无法或难以查找的邮件充塞管理机关的监控存储，而管理机关又无法追查邮件来源，或者追查邮件来源时需要消耗很大的精力，无法保证立法的顺利实施，损害了法律的威严。

同时，对于国外的、管理机关无法监管的计算机发出的垃圾邮件，如果不基于可追查检查，则黑名单方案很难奏效，可能导致殃及无辜，而且垃圾邮件发送者可以躲避黑名单方案。而使用了可追查检查方案后，对于正常的邮件服务器和邮件发件人，如果因为当地没有使用全面可追查方案而被误判，可以从错误反馈信息中得到原因，使用另外可追查方案可接受的邮件发送方式和国内使用了可追查检查方案的系统通邮。如果国外计算机使用可追查检查方案可接受的方式发送垃圾邮件，黑名单方案就可以起作用了。

立法以后无法实施、难以实施，对违法的人难以查找和制裁，必然损害法律的威严。

2. 可追查后的立法实施

使用了可追查性检查后，所有能够在网络上传输的邮件都是可以追查来源的，这时内容监控也可以起作用了，一方面可追查体系保证了邮件发件人的可追查性，一般人慑于法律威严不敢乱发垃圾邮件；另一方面内容监控提交的信息都是可追查到的，管理机关可以很容易地有的放矢，可以很容易地找到和制裁需要制裁的邮件发件人。

立法的可实施性和易实施性，提高了管理机关的工作效率，维护了法律的威严。