2016 年 8 月 2 日，星期二

来自宏的威胁：悄无声息地越过 Office 的防线

本博文由 Matthew Molyett 撰写。特别感谢 Martin Lee 提供的建议。

引言

早在 20 世纪 90 年代中期，宏就成为散播恶意软件和感染系统的工具。从 90 年代末到 21 世纪初，随着越来越多的用户意识到需要禁用 Microsoft Word 的宏功能，基于宏的恶意软件也遭到了遏制。但是，Microsoft (MS) Office 文件格式在 2007 年发生了变化，为隐藏宏的存在提供了可乘之机，所以利用宏传播恶意软件的行为又呈现出增长之势。

本文将探讨 MS Office 文件格式如何被用于恶意目的和迷惑用户，以及宏恶意软件的传播范围。

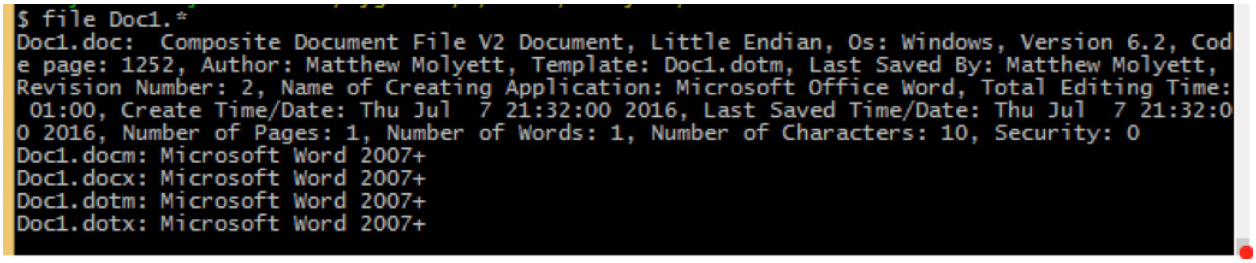


图 1：五种不同的 Microsoft Word 文件所使用的文件实用程序

文档文件中的“宏”是什么概念？

文档与宏

Microsoft Office 提供了名为 Visual Basic for Applications 的编程语言。作为宏语言，它功能齐全，可以嵌入到文件中，实现任务自动化。自我传播型病毒（例如 20 世纪 90 年代末的“美丽杀手”[Melissa]）就是利用宏功能默认执行行为的特性大肆传播。

MS Office 2003 默认禁止执行宏，而且如果文件中存在宏，图形用户界面 (GUI) 会弹出提示消息通知用户，从而抑制了宏的自动执行特性。MS Office 2007 在宏保护方面又迈进了一大步：默认 MS Word 文档文件格式不支持宏。为了做到这一点，Microsoft 基于 OfficeOpen XML 标准引入了 4 种不同的文件格式：

文件扩展名	文件类型	是否允许使用宏
DOCX	压缩文档	不支持
DOTX	压缩模板	不支持
DOCM	压缩文档	支持
DOTM	压缩模板	支持

与通过检测文件内容确定文件类型的基于 Unix 的操作系统不同，MS Windows 以文件扩展名（即文件名中最后一个“.”之后的字符）为基础，决定文件被点击后系统使用哪个应用打开文件。MS Office 安装完毕后，会自动与上面列出的扩展名相关联。因此，当您点击上面列出的所有文件类型时，系统都会使用 MS Word 打开文件。

DOCX - 不允许使用宏！

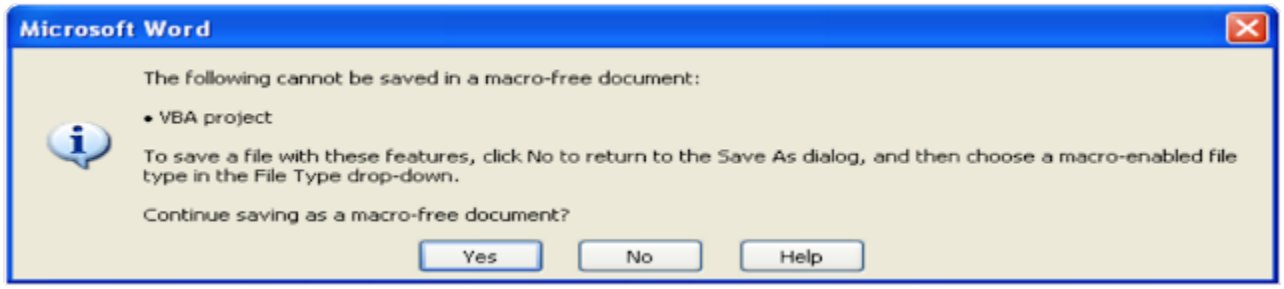


图 2：尝试将宏代码保存到 DOCX 文件时显示的消息

DOC 文件（MS Office 2007 以前的 MS Word 版本所使用的文件格式）允许在文档中嵌入多种组件，包括宏。用户在打开文件之前无法确定文档是否安全。MS Office 2007 中集成的 OfficeOpen XML (OOXML) 标准消除了这种顾虑。上面的每种文件格式都是 zip 存档文件，其中包含通用格式的 XML 文件。

该存档文件中的 [Content_Types].xml 组件会为文件中的其他组件提供 MIME 类型的信息。MS Word 支持的这四种文件格式都具有独特的 MIME 类型。仅当 MIME 类型与 DOCM 和 DOTM 文件格式关联时，才可以保存或运行宏。如果 Content_Types 组件断言 MIME 类型为 DOCX 或 DOTX，MS Word 将不会保存或运行宏代码。

要是将 DOCX 重命名为 DOCM，可以添加宏代码吗？

自然有人会问，如果将 DOCX 文件重命名为 DOCM，是不是就能向文件中添加宏了呢？其实，系统会通过检查 OOXML 文件格式来分析文件扩展名（MIME 类型协议），所以答案是“不行”。

Microsoft Word 在打开文档时，首先会检查文件名，以确认文档是否为 OOXML 文件。如果打开错误的 DOCM 文件，系统会发现 MIME 类型与文件数据中包含的 DOCX 信息不符，并弹出错误消息。

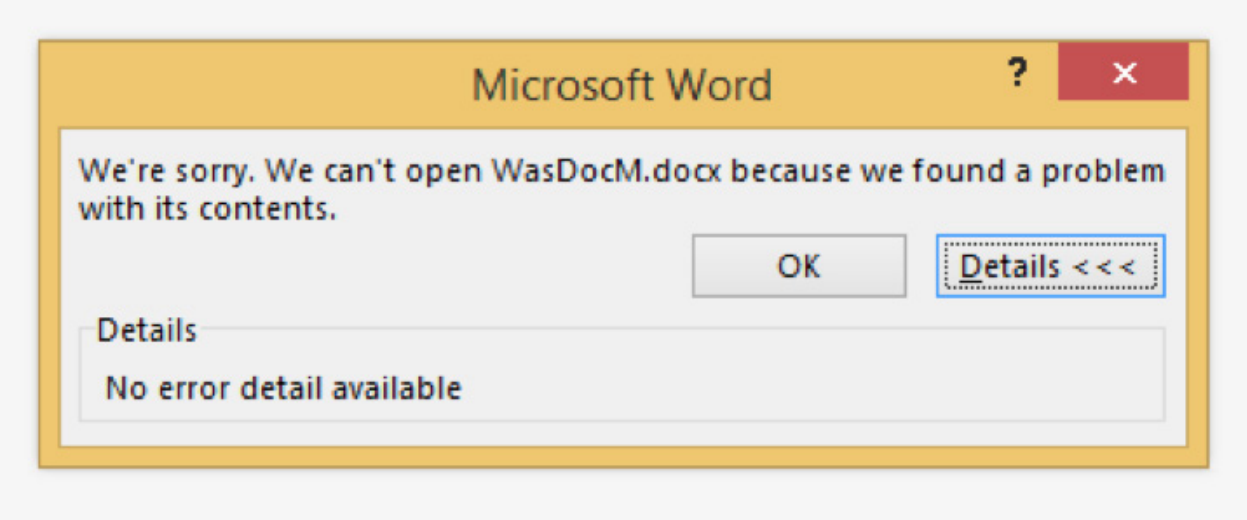


图 3：重命名的 DOCM 文件仍会以 DOCX 文件名打开

文件扩展名	MIME 类型
DOCX	*application/vnd.openxmlformats-officedocument.wordprocessingml.document.main+xml*
DOCM	*application/vnd.ms-word.document.macroEnabled.main+xml*
DOTX	*application/vnd.openxmlformats-officedocument.wordprocessingml.template.main+xml*
DOTM	*application/vnd.ms-word.template.macroEnabledTemplate.main+xml*

包含宏的 OOXML 文档是不是必须命名为 DOCM？

一般而言，MS Word 会根据文件数据（而不是文件扩展名）打开文件。只要 MS Word 能够识别文件的数据结构，就能正确打开文件。如果系统将某个文件识别为 MS Office 2007 文件，该文件的内部必须具有正确的 MIME 类型，否则会造成验证错误，导致文件无法打开。

OOXML 文件类型由 MS Office 的 WWLIB.DLL 组件进行验证，该组件会确认文件的 MIME 类型是否与预期一致。如果文件扩展名未表明文件是 OOXML 文件类型，这步验证操作通常会直接通过（即使 MIME 类型确实是 OOXML）。这意味着如果包含宏的 OOXML 文档（DOCM 或 DOTM 文档）使用其他文件扩展名，就能成功实现加载。也就是说，即使 OOXML 文件使用非 OOXML 文件扩展名，只要采用 MS Word 已注册的处理格式，就能成功打开。

因此，包含嵌入宏的 DOCM 文件可以通过改变文件扩展名伪装成其他文件格式。例如，RTF 文件格式不支持 MS Office 宏代码，但是如果将 DOCM 文件重命名为 RTF 文件，就能在 MS Office 中打开，使嵌入的宏代码成功运行。这种策略目前已普遍用于漏洞攻击。

图 4：WWLIB.DLL 验证 MIME 类型是否为 DOCX

默认文件数据识别和 OOXML

2016 年 5 月，我们开始发现我们收到的一些样本虽然扩展名是 DOC、RTF 和 DOT，但实际上 MIME 类型却是 *application/vnd.ms-word.template.macroEnabledTemplate.main+xml*（即 DOTM）。默认情况下，程序通过分析文件类型得到的结果会是“Microsoft Word 2007+”（见图 1），所以系统会将这些文件作为 DOCX 文件来执行。这些样本并未展示出恶意行为，仅仅引发了错误提示（类似于图 3 所示的错误消息）。

这意味着什么？

这类攻击已经普遍存在

Talos 对这些支持宏的模板（在本部分特指 DOTM）文件的出现情况进行了持续跟踪，发现这些文件的部署速度在过去几个月内一直呈现快速增长趋势。我们收集了 3 月 18 日至 7 月 13 日期间发现的所有 DOTM 文件，并对宏负载进行了检查。分析发现，许多文档中都重复使用了一种欺骗计算机的宏。

在发现发生冲突后，我们重点收集了至少在 4 个不同的 DOTM 文件中出现的宏冲突，以进行进一步检查。在这四个月发现的所有 DOTM 文件中，存在问题的 DOTM 文件多达 64%。

收集方法

在这项分析中，我们使用 VirusTotal 的文件库作为 DOTM 文件的来源。DOTM 文件可通过 MIME 类型字符串“application/vnd.ms-word.template.macroEnabledTemplate”进行识别，所有选定文件均以“首次提交日期”为准。

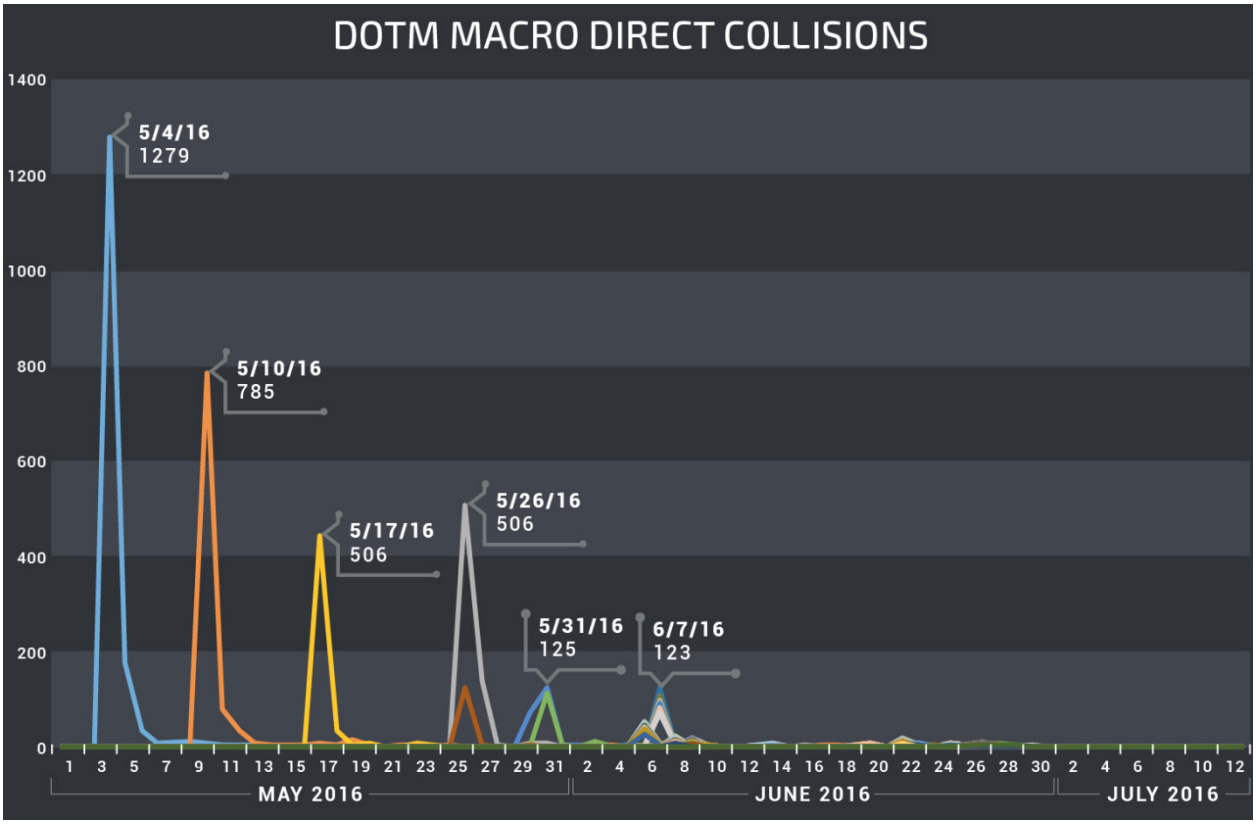
隔离方法

我们使用 ClamAV 提供的辅助命令行组件“sigtool”来加快签名的生成速度，该组件可以从 Microsoft Office 文档文件中提取宏文本。为了识别宏，我们解压了所有 DOTM 存档文件，找出其中包括的 ./word/vbaProject.bin 文件。如果使用 sigtool 执行提取后在不同的 vbaProject.bin 文件中得到相同的代码片段，就可以确定宏是重复的。

在收集到的所有宏中，有 5% 的宏存在于 64% 的 DOTM 文件中。在最主要的 255 个冲突中，我们发现其中一个是空文件：不包含任何 vbaProject.bin 的 DOTM 文件。我们没有将该文件纳入分析中，所以我们研究的冲突数量为 254 个。

表 2：DOTM 和宏文件的数量

宏总数	6054
DOTM 文件总数	16377
唯一宏	5641
罕见宏（出现不超过 4 次）	99
常见宏（出现 4 次或 4 次以上）	314
包含常见宏的 DOTM	10518



图表 1：直接的宏冲突（254 个不同样本）

分组方法

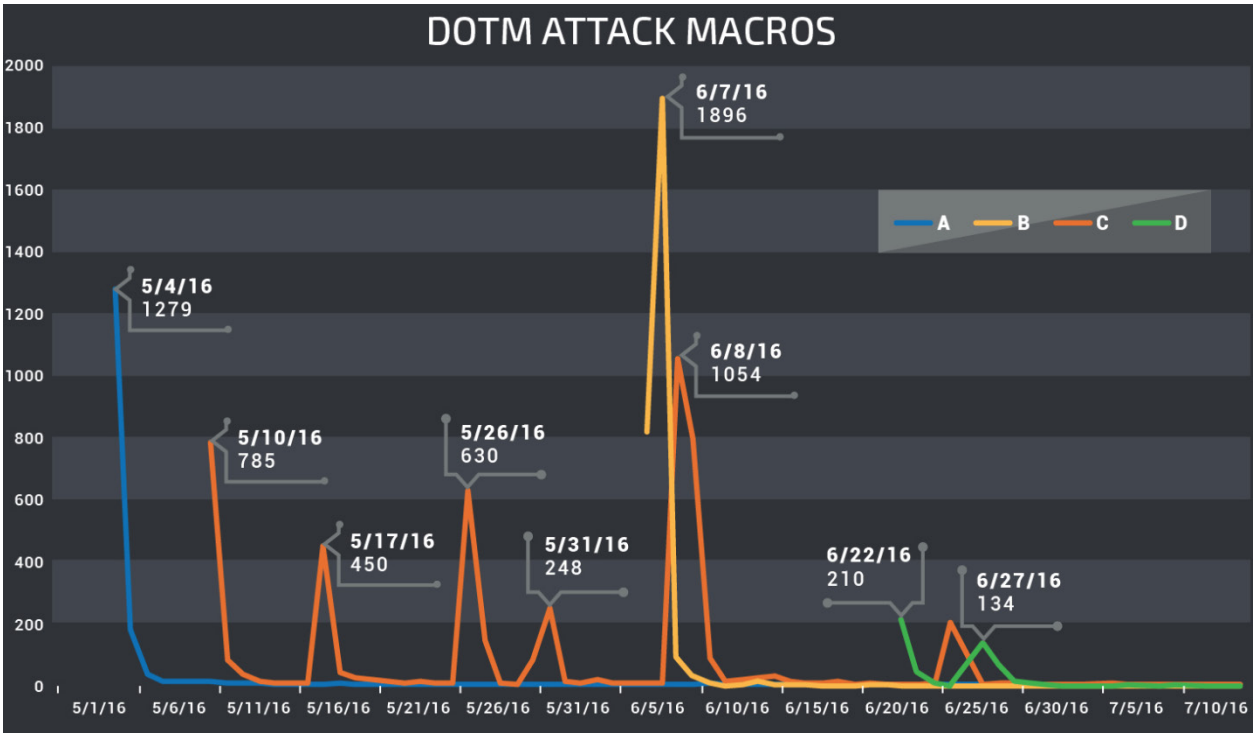
从所确定的宏冲突可以明显看出，整个五月发生了多次攻击活动。虽然在图表 1 中，六月和七月未显示出高峰值，但是收集到的 DOTM 文件数量也超过了 5000 个。我们认为，在五月，原有的伪装方法或变化方法有所改进，或者有更老道的攻击者加入了这个领域。为了进行进一步分组，我们使用了 ClamAV 扫描引擎。

我们生成了 19 种不同的 LDB 签名，对 254 个样本进行分组。其中有 9 种签名是通过自动分析宏脚本中使用的字符串生成的，另外 9 种签名是手动创建的，用于区别同一种伪装方法的多个形式。在这 18 种签名中，除了 3 个主要冲突外，其他的签名都能得到识别。我们对这 3 个冲突中的 2 个进行了检查，以确定它们是否为普通的宏，结果发现其中一个宏是空的。但是，另一个冲突却采用了与其他 18 种签名完全不同的独特伪装方法，于是我们专门为它手动创建了一个签名。

根据签名，我们将这些宏分为 4 个不同的组。有多个签名出现在多个组中，这些签名对区分宏帮助不大。用于检测文件关闭时触发宏的签名分为一组；用于创建直接数组的签名分为一组；用于创建复杂数组的签名分为一组；用于检测字符串中嵌入的共享数据的签名分为一组。

```
Private Function CRE4C(ByVal As47kS As String, ByVal O1WA As Long, ByVal TwPCt As Variant) As String
Dim OqKPo3j() As Byte, DT5() As Byte, XGcj2 As Long, E1O0 As Long
OqKPo3j = As47kS
XGcj2 = IBEL(OqKPo3j)
O1WA = (O1WA - 1) * 2
TwPCt = (TwPCt * 2) - 1
If O1WA + TwPCt > XGcj2 Then TwPCt = XGcj2 - O1WA
ReDim DT5(TwPCt)
For E1O0 = O1WA To O1WA + TwPCt
DT5(E1O0 - O1WA) = OqKPo3j(E1O0)
Next E1O0
CRE4C = DT5
End Function
```

图 5：明显的由计算机生成的宏，只有 8 个冲突



图表 2：攻击宏分组（4 组，分别标记为 A、B、C 和 D）

需要研究的其他领域

本文主要侧重于 MS Word，但是 MS Excel 和 PowerPoint 也存在类似的 OOXML 格式。带有嵌入宏的 PPTM 文件可以伪装成无害的 PPT 演示文稿。在更糟的情况下，通过相同的方法，攻击者可以将带有嵌入宏的 MS Excel XLSM 文件伪装成 CSV 纯文本电子表格文件，让 Excel 毫无防备地打开文件，运行其中包括的代码。

对抗宏攻击的方法

您可以安装 WWLIB 验证功能的相应补丁，以便在遇到 MIME 类型为 DOCM 或 DOTM 的文件时验证文件扩展名是否与预期一致，从而轻松检测并阻止攻击。

支持宏的模板是一种极少使用和传输的文件类型。在网络网关阻止文件格式传入可能不会违反商业惯例，而且可以有效地让恶意文件远离终端用户。

总结

自 2003 年以来，Microsoft Word 加入了许多保护措施来防止执行文档中嵌入的恶意宏。MS Office 2007 文件格式对可以包含宏的文件类型和不可以包含宏的文件类型进行了区分。但是，这种依靠文件类型的保护措施并不充分。威胁发起者已经发现，只要将包含嵌入宏的 OOXML 文档伪装成其他文件类型，就能成功躲过 Microsoft Office 的文件类型检测。在 MS Office 的验证机制得到完善之前，用户应时刻防范预期之外的 MS Office 文档，因为即使是“安全的”文件格式，也仍有可能包含恶意代码。

检测依据

思科客户可以通过下列途径检测和阻止宏威胁：

产品	保护
AMP	✔
CWS	✔
ESA	✔
网络安全	✖
WSA	✔

- 高级恶意软件防护 (AMP) 可以有效防止恶意软件执行。
- CWS 或 WSA 的 Web 扫描功能可以阻止访问恶意网站，并检测恶意软件。
- ESA 可以拦截威胁发起者在攻击活动中发出的恶意电子邮件。
- ClamAV 可以通过 *OXmlEvader* 的形式检测宏攻击。

发布者：MARTIN LEE 发布时间：9:42
标签：宏、恶意软件、OFFICE、WORD 文档