

MSHTML [Let's defend – Write-up]

2021's 0-Day MSHTML

file location: /root/Desktop/ChallengeFiles/Employee_W2_Form.docx

file location2: /root/Desktop/ChallengeFiles/Employees_Contact_Audit_Oct_2021.docx

file location3: /root/Desktop/ChallengeFiles/Work_From_Home_Survey.doc

file location4: /root/Desktop/ChallengeFiles/income_tax_and_benefit_return_2021.docx

| Start Investigation

```
root@ip-172-31-3-35:~/Desktop/Tools# python3 zipdump.py /root/Desktop/ChallengeFiles/Employees_Contact_Audit_Oct_2021.docx
Index  Filename                               Encrypted  Timestamp
1  [Content_Types].xml                       0  1980-01-01 08:00:00
2  _rels/                                     0  2021-09-01 23:23:30
3  _rels/.rels                               0  1980-01-01 08:00:00
4  docProps/                                 0  2021-09-01 23:23:30
5  docProps/app.xml                          0  1980-01-01 08:00:00
6  docProps/core.xml                         0  1980-01-01 08:00:00
7  word/                                     0  2021-09-01 23:23:30
8  word/_rels/                               0  2021-09-01 23:23:30
9  word/_rels/document.xml.rels              0  2021-09-10 16:37:10
10 word/document.xml                         0  2021-09-01 23:23:30
11 word/fontTable.xml                       0  1980-01-01 08:00:00
12 word/media/                              0  2021-09-01 23:23:30
13 word/media/image1.wmf                    0  1980-01-01 08:00:00
14 word/settings.xml                        0  1980-01-01 08:00:00
15 word/styles.xml                          0  1980-01-01 08:00:00
16 word/theme/                              0  2021-09-01 23:23:30
17 word/theme/theme1.xml                    0  1980-01-01 08:00:00
18 word/webSettings.xml                     0  1980-01-01 08:00:00
root@ip-172-31-3-35:~/Desktop/Tools#
```

To enhance our analysis, we can pipe the output into another powerful tool developed by Didier Stevens: re-search.py.

re-search.py is a specialized utility that leverages regular expressions (regex) to parse and search through data. It allows you to utilize a built-in library of common patterns or define your own custom expressions for more targeted searches.

By combining zipdump and re-search, we can execute a command that dumps all file indexes from the sample and pipes them directly into the search tool. We will then apply specific filters to automatically identify and extract all unique IPv4 addresses from the output.

Q1: Examining the Employees_Contact_Audit_Oct_2021.docx file, what is the malicious IP in the docx file?

Analyzing OOXML Files with zipdump.py

Since .docx files are essentially compressed ZIP archives, we can explore their internal structure using zipdump.py (by Didier Stevens), as recommended in the SANS cheat sheet.

To examine the contents of the document, run the following command:

```
python3 zipdump.py
/root/Desktop/ChallengeFiles/Employees_Contact_Audit_Oct_2021.docx
```

```
root@ip-172-31-3-35:~/Desktop/Tools# python3 zipdump.py -D /root/Desktop/ChallengeFiles/Employees_Contact_Audit_Oct_2021.docx | python3 re-search.py -n -u ipv4
175.24.190.249
root@ip-172-31-3-35:~/Desktop/Tools#
```

Answer: 175.24.190.249.

Q2: Examining the Employee_W2_Form.docx file, what is the malicious domain in the docx file?

To answer the question, I extracted the file hash and searched in Virustotal.

```
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# ls
Employee_W2_Form.docx      Work_From_Home_Survey.doc
Employees_Contact_Audit_Oct_2021.docx  income_tax_and_benefit_return_2021.docx
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# md5sum Employee_W2_Form.docx
45e7d6562bfddb816d45649dd667abde Employee_W2_Form.docx
root@ip-172-31-3-35:~/Desktop/ChallengeFiles#
```

679bbe0c50754853978a3a583505ebb99bce720cf26a6aaf8be06cd879701ff1

Joe Sandbox Analysis:

Verdict: MAL
Score: 64/100
Threat Name: CVE-2021-40444
Domains: prod-wus-resolver.naturallanguageeditorservice.osi.office.net, arsenal.30cm.tw, arsenal.30cm.tw
Hosts: 52.110.2.130, 72.154.7.38
HTML Report: <https://www.joesandbox.com/analysis/1857440/0/html>
PDF Report: <https://www.joesandbox.com/analysis/1857440/0/pdf>
Executive Report: <https://www.joesandbox.com/analysis/1857440/0/executive>
[Show more](#)

FileScanIO
6 months ago

FileScanIO Analysis:

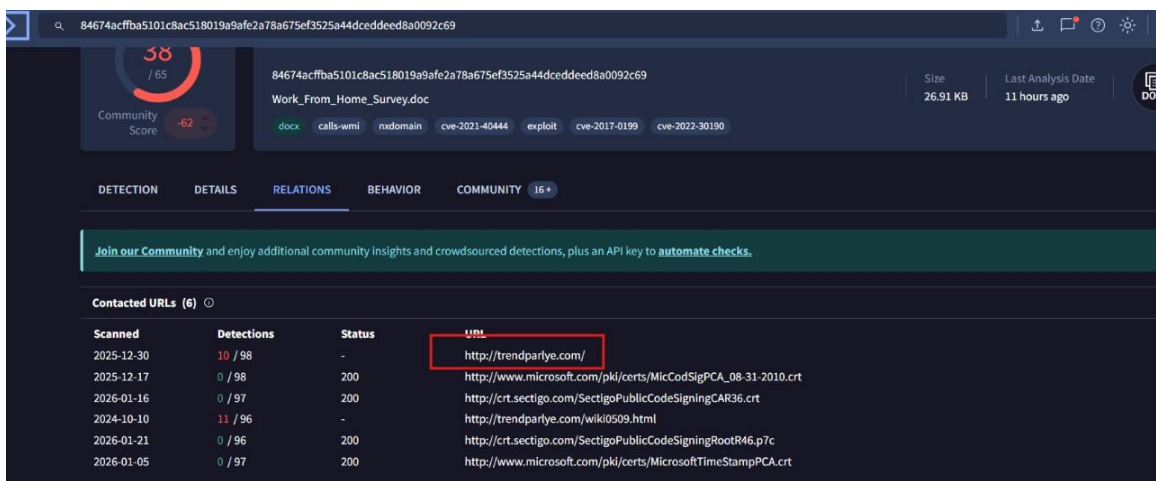
Verdict: MALICIOUS
Confidence: 100/100
Tags: docx, ooxml, arsenal.30cm.tw
Domains: arsenal.30cm.tw, arsenal.30cm.tw
Hosts: 128.199.107.104, 128.199.107.104
Report: <https://www.filescan.io/reports/679bbe0c50754853978a3a583505ebb99bce720cf26a6aaf8be06cd879701ff1/e005d7c3-4b01-4456-bb5d-38af2b8a4a66>

Answer: arsenal.30cm.tw

Q3: Examining the Work_From_Home_Survey.doc file, what is the malicious domain in the doc file?

Also to answer this question, I extracted the file hash and searched in Virustotal.

```
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# ls
Employee_W2_Form.docx      Work_From_Home_Survey.doc
Employees_Contact_Audit_Oct_2021.docx  income_tax_and_benefit_return_2021.docx
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# md5sum Employee_W2_Form.docx
45e7d6562bfddb816d45649dd667abde Employee_W2_Form.docx
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# md5sum Work_From_Home_Survey.doc
41daca2a33ee717abcc8011b705f2cb Work_From_Home_Survey.doc
root@ip-172-31-3-35:~/Desktop/ChallengeFiles#
```



The screenshot shows the VirusTotal interface for the file 'Work_From_Home_Survey.doc'. The file has a Community Score of 38/65 and a detection rate of 62%. It is categorized as a document (docx) and is associated with several CVEs: CVE-2021-40444, CVE-2017-0199, and CVE-2022-30190. The 'Contacted URLs' section is expanded, showing a list of URLs scanned by various engines. The URL 'http://trendparlye.com/' is highlighted with a red box, indicating it was detected by Trend Micro.

Scanned	Detections	Status	URL
2025-12-30	10 / 98	-	http://trendparlye.com/
2025-12-17	0 / 98	200	http://www.microsoft.com/pki/certs/MicCodSigPCA_08-31-2010.crt
2026-01-16	0 / 97	200	http://crt.sectigo.com/SectigoPublicCodeSigningCAR36.crt
2024-10-10	11 / 96	-	http://trendparlye.com/wiki0509.html
2026-01-21	0 / 96	200	http://crt.sectigo.com/SectigoPublicCodeSigningRootR46.p7c
2026-01-05	0 / 97	200	http://www.microsoft.com/pki/certs/MicrosoftTimeStampPCA.crt

Answer: trendparlye.com

Q4: Examining the income_tax_and_benefit_return_2021.docx, what is the malicious domain in the docx file?

Also to answer this question, I extracted the file hash and searched in Virustotal.

```
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# ls
Employee_W2_Form.docx      Work_From_Home_Survey.doc
Employees_Contact_Audit_Oct_2021.docx  income_tax_and_benefit_return_2021.docx
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# md5sum Employee_W2_Form.docx
45e7d6562bfddb816d45649dd667abde Employee_W2_Form.docx
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# md5sum Work_From_Home_Survey.doc
41daca2a33ee717abcc8011b705f2cb Work_From_Home_Survey.doc
root@ip-172-31-3-35:~/Desktop/ChallengeFiles# md5sum income_tax_and_benefit_return_2021.docx
55998cb43459159a5ed4511f00ff3fc8 income_tax_and_benefit_return_2021.docx
md5sum: q: No such file or directory
root@ip-172-31-3-35:~/Desktop/ChallengeFiles#
```

File: income_tax_and_benefit_return_2021.docx
Size: 23.61 KB
Last Analysis Date: 4 days ago
Community Score: -36

CONTACTED URLS (2)

Scanned	Detections	Status	URL
2026-01-16	13 / 97	-	http://hidusi.com/8c76295a59acb7/side.html
2026-01-30	0 / 94	405	https://mobile.pipe.aria.microsoft.com/Collector/3.0/

Answer: hidusi.com

Q5: What is the vulnerability the above files exploited?

The answer is also found in VirusTotal.

File: income_tax_and_benefit_return_2021.docx
Size: 23.61 KB
Last Analysis Date: 4 days ago
Community Score: -36

45/67 security vendors flagged this file as malicious

Popular threat label: trojan.w97m/cve202140444

Threat categories: trojan, downloader

Family labels: w97m, cve202140444, expl

Security vendors' analysis

Vendor	Detection	Category
AhnLab-V3	Downloader.DOC.External	Alibaba
Alibaba	Trojan:Office.CVE-2021-40444	Alibaba
AliCloud	Exploit:MSOffice/CVE-2021-40444.XML	ALYac
Avast	Trojan.Downloader.DOC.Gen	Avast
Arcabit	Trojan.Generic.D2CB227F	Avast
Avast	XML.CVE-2021-40444-A [Exploit]	Avast

Answer: cve-2021-40444

Summary

By analyzing multiple samples like Employee_W2_Form.docx and Work_From_Home_Survey.doc, I learned how to use zipdump.py and re-search.py to peel back the layers of OOXML structures and extract hidden malicious IPs and domains. This lab was particularly beneficial because it demonstrated the "real-world" transition from a simple document to a network-based attack, showing me how to pivot from local file analysis to global threat intelligence tools like VirusTotal to identify specific vulnerabilities.