## Eagle Eye

샌드박스 시대의 종결, AI 에뮬레이터 시대의 도래



최원혁 : ㈜누리랩 대표/연구소장

Eagle Eye: 2008년 개봉한 영화로 AI 시스템 이글아이가 핸드폰, 현금지급기, 거리의 CCTV, 교통안내 LED사인보드, 신호등 등 우리 주변의 전자장치와 시스템을 모니터링 하여 사람들을 AI가 원하는 방향으로 행동하도록 조정하는데에서 착안

Code # Engn

www.CodeEngn.com
2021 CodeEngn Conference 17

CGmodeling.con

# 이력 활동

#### 최원혁

hanul93@gmail.com



1993.07: 악성코드 분석 시작

1995.10: 키콤백신 개발 발표(국내 4번째 백신)

1998.03: ㈜하우리 창업(악성코드 분석 및 바이로봇 개발 총괄/CTO)

2009.03: 동국대학교 국제정보대학원 정보보호학과 사이버포렌식전공 교수

2015.05 : ㈜누리랩 창업(대표/연구소장)

2020.09: 이화여자대학교 엘텍공과대학 소프트웨어학부 사이버보안 겸임교수

2021.03: 고려대학교 인공지능사이버보안학과 겸임교수

2004.08: 민관합동조사단 1~4기 활동 (악성코드 분석 전문)

2009.01: 국가정보원장 표창 수상

2014.03: 사이버보안전문단 (미래창조과학부 임명)

2021.03: 사이버작전사령부 자문 위원

2017.09: 파이썬으로 배우는 Anti-Virus 구조와 원리

•

#### 인터넷, 고객사를 통해 모은 악성코드 샘플들…

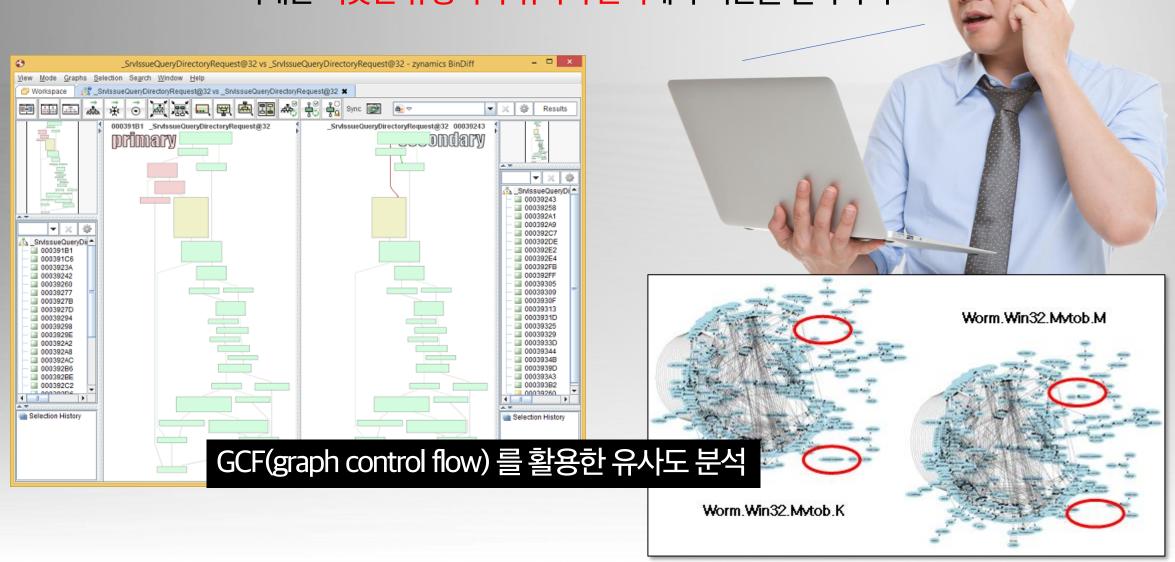
#### 많아도 너~무 많아!!!





#### 유사도의 필요성 대두

최대한 비슷한 유형끼리 묶어서 분석해야 시간을 절약하지

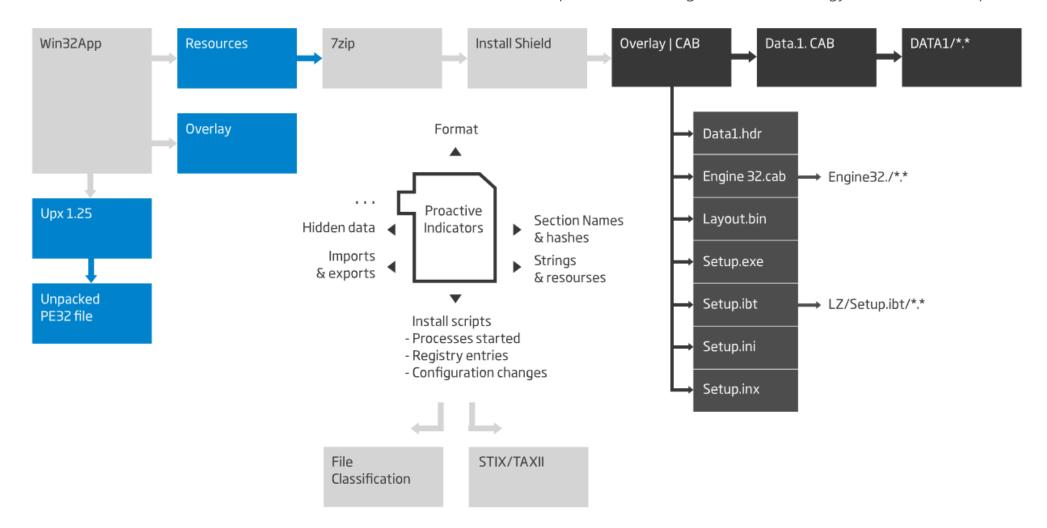






#### 언팩을 해야 하는데…

이미지 출처: https://www.reversinglabs.com/technology/active-file-decomposition



샌드박스를 사용하여 언팩을 진행하면 알려지지 않은 패킹도 처리할 수 있습니다.

#### → 하지만, 샌드박스 우회 악성코드도 있죠!!

→ 그래서 이런 연구도 있습니다.

△ 그걸 또 우회하는 것도 있죠!



#### Cuckoo Sandbox를 회피하는 악성코드 탐지 방안 연구

정제성\*, 김광조\*

\*한국과학기술원 전산학부

A Study on Detection of Evasive Malware in Cuckoo Sandbox

Je-Seong Jeong\*, Kwangjo Kim

\*School of Computing, KAIST

#### 요약

악성코드는 정상적으로 시스템 사용을 못하게 하는 것으로서 컴퓨터 바이러스 부터 웹, 트로이 목마 등 그 종류도 다양하다. 이러한 악성코드를 분석하는 방법에는 정정분석과 동적분석이 있지만 요즘과 같이 악성코드의 수가 많은 시대에는 코드를 분석하는 정적분석 보다는 악성코드를 실색시켜 행위를 분석하는 동적분석이 시간적인 측면에서 봤을 시 좀 더 효율적이다. 그리고 요즘 동적분석으로 많이 사용하고 있는 것 중의 하나는 오픈소스 악성코드 자동 분석 시스템인 쿠쿠 센드박스(Cuckoo Sandbox)이다. 하지만 요즘 약성코드 분석 시스템을 우회하는 기능까지 탑재하여 분석을 어렵게 하고 있다. 따라서 본 논문 에서는 이런 우회 기법을 분석 후 쿠쿠 센드박스(의 적용하여 탐지 기능을 향상시키는 방안을 제안하고자 한다.

#### 악성코드를 무조건 실행할 수 있는 방법이 없을 까?



② 악성코드가 특정 조건에만 동작하는 것도 무시하고 싶다. (1월 1일에 파괴 동작하는 악성코드, 포르투갈 윈도우에서만 동작하는 악성코드 등)

③ 지금은 동작하지 않는 C&C라도 이 악성코드 분석은 계속 하길 원해…

④ 그냥 다 분석하고 싶어~~~~~~



# 에뮬레이터

#### 기존에뮬레이터… 운영체제를 예뮬레이션 할 수 없다.

#### WINE은 리눅스에서 윈도우용 프로그램을 에뮬레이션한다.

즉, 윈도우 API를 적절하게 리눅스 API에 맵핑하면 된다.

아하!!그럼… 실행파일로더와 API를 적절하게 맵핑하면 되겠군.



#### 세상은 넓고 나와 같은 생각을 하는 사람은 많다.



이미 이런 프로젝트가 존재합니다.

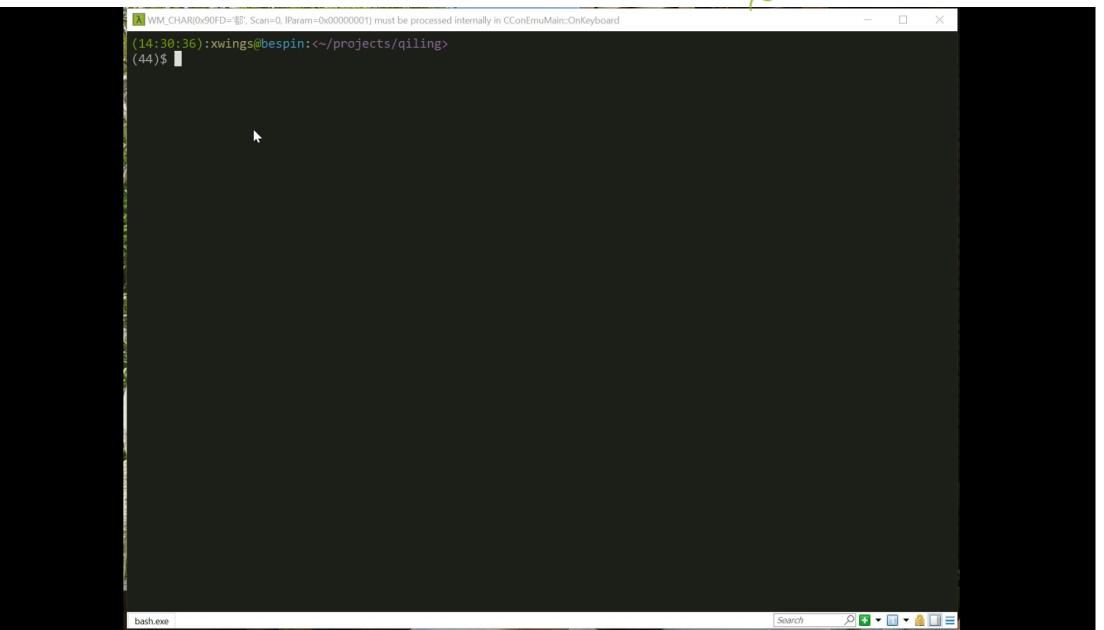
















#### 앗!!! 모든 윈도우 API를 에뮬레이션 못하네?

코딩 되지 않은 윈도우 API는 실행하지 못함



즉, 악성코드 실행을 위해 직접 필요한 API를 모두 코딩 해야 함





# 에뮬레이터



#### 다양한 운영체제 환경을 에뮬레이션 가능성 확인

API의 리턴 값은 하드 코딩된 값 만을 리턴



dllname = 'kernel32\_dll'

# BOOL GetModuleInformation(
# HANDLE hProcess,
# HMODULE hModule,
# LPMODULEINFO lpmodinfo,
# DWORD cb
# );
@winsdkapi(cc=STDCALL, dllname=dllname)
def hook\_K32GetModuleInformation(self, address, params):
 # TODO
 return 0

특정일에 동작하는 악성코드의 경우, 날짜 API를 호출하여 얻은 결과를 통해 **악성행위**를 할지를 결정

**하드 코딩된 API 리턴 값**은 의미 없음



운영체제를 에뮬레이션 하고…

악성코드가 원하는 API 리턴 값을 잘 전달하며

최소한의 API 코딩으로 악성코드 행위를 추적할 수 있다면…

**샌드박스**로부터 자유를 외칠 수 있다.





#### 분석 대상 프로그램 구동은 단순하게…

```
# Author: Kei Choi (hanul93@gmail.com)
# NEmu 예제 프로그램

from nemu.nemu import NEmu

pif __name__ == '__main__':
    ne = NEmu(['sample/abexcml.exe'], verbose=True) # NEmu 준비 ne.run() # 파일 실행
```

플러그인 방식으로 대상 파일 포맷 분석····

대상 파일에 맞는 **운영체제 초기화** 

운영체제 메모리에 파일 로딩 후 실행

```
13
        def run(self): # 에뮬레이터를 실행
14
            # 파일 분석 (플러그인 방식으로)
15
            filename = self. run args[0]
16
17 E
            plugins = [
                nefile pe.NEmuFile(verbose=self. verbose),
18
19
20
21
            for plugin in plugins:
22
                if plugin.format(filename):
23
                     break
24
            else:
25
                plugin = None
26
            # TODO 파일 포맷에 맞는 OS 부팅(?)
            if not plugin:
29
                 return False
30
            else:
31
                plugin.os init()
32
33
             # TODO 파일 실행
34
            plugin.run()
```



#### 에뮬레이터

```
def run(self): # 에뮬레이터를 실행
14
            # 파일 분석 (플러그인 방식으로)
            filename = self. run args[0]
16
17
            plugins = [
18
                nefile pe.NEmuFile(verbose=self. verbose)
19
            for plugin in plugins:
                if plugin.format(filename):
23
                    break
24
            else:
                plugin = None
26
            # TODO 파일 포맷에 맞는 OS 부
            if not plugin:
29
                return False
30
            else:
                plugin.os init
32
33
34
            plugin.run()
```

#### 프로그램 실행

Capstone 사용하여 에뮬레이션

```
C:\WINDOWS\system32\cmd.exe
                                                                                                                                                     _ _
      운영체제 초기화
                                                         Windows(32bit) is Booting
Copyright 2015–2021 Nurilab Inc.
[+] Loding FileName: sample/abexcm1.exe
[+] File Image Base: 0x00400000
   0x70000000, kernel32.dll:GetDriveTypeA
0x70000020, kernel32.dll:ExitProcess
0x70000040, user32.dll:MessageBoxA
00401000: 6a 00
00401002: 68 00204000
00401007: 68 12204000
                                                                0x402000
                                                                                                            ; "abex' 1st crackme"
; "Make me think your HD is a CD-Rom."
                                                       push
                                                                0x402012
                                                       push
                                                       push
0040100e: e8 4e000000
                                                                0×401061
00401061: ff25 5c304000
                                                                dword ptr [0x40305c]
00401013: 68 94204000
00401018: e8 38000000
                                                                                                            ; "c:\"
                                                                0×402094
                                                                0x401055
  dword ptr [0x403050]
00401055: ff25 50304000
0040101d: 46
0040101e: 48
                                                                esi
                                                                eax
0040101f: eb 00
                                                                0x401021
```



#### <del>- 운영체제를 에뮬레이션 하고····</del> ≪

악성코드가 원하는 API 리턴 값을 잘 전달하며

최소한의 API 코딩으로 **악성코드 행위를 추적할 수 있다면**…

**샌드박스**로부터 자유를 외칠 수 있다.





#### 악성코드가 원하는 API 리턴 값

특정일에 동작하는 악성코드의 경우, 날짜 API를 호출하여 얻은 결과를 통해 **악성행위**를 할지를 결정

```
00401055: ff25 50304000
                                                  dword ptr [0x403050]
                                                                                   : kernel32.dll:GetDriveTvpeA
                                          imp
                                                                                              윈도우 API 호출 이후
0040101d: 46
                                          inc
                                                  esi
                                          dec
                                                  eax
                                                  0 \times 401021
                                           jmp
                                          inc
                                                  esi
                                          inc
                                                  esi
                                          dec
                                                  eax
                                                                                   cmp 중 eax와 비교 되는 값…
                                                  eax, esi
                                           CMD
                                                  0x40103d
                                          jе
                                          push
                                                                                     "Error"
"Nah... This is not a CD-ROM Drive!"
   0102a: 68 35204000
                                                  0x402035
                                          push
         68 3b204000
                                                  0x40203b
                                          naud
                                          push
00401036: e8 26000000
                                                  0x401061
                                          call
```

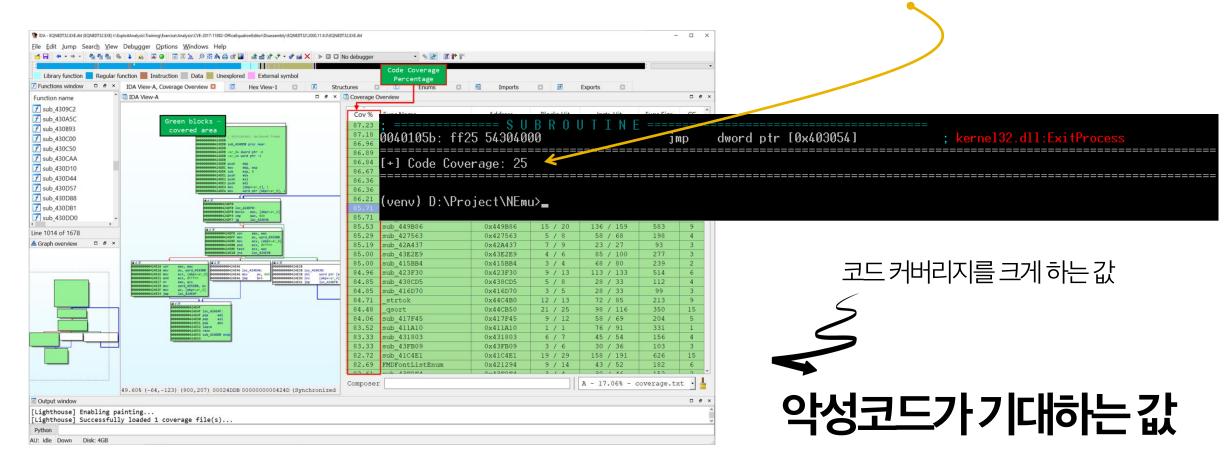
윈도우 API 기대 값



#### 에뮬레이터

#### 윈도우 API 기대 값에 따라 분기 발생

악성코드가 원하는 기대 값인가 아닌가는 **코드거버리지**로 측정







나는 1월 1일에 하드디스크 파괴



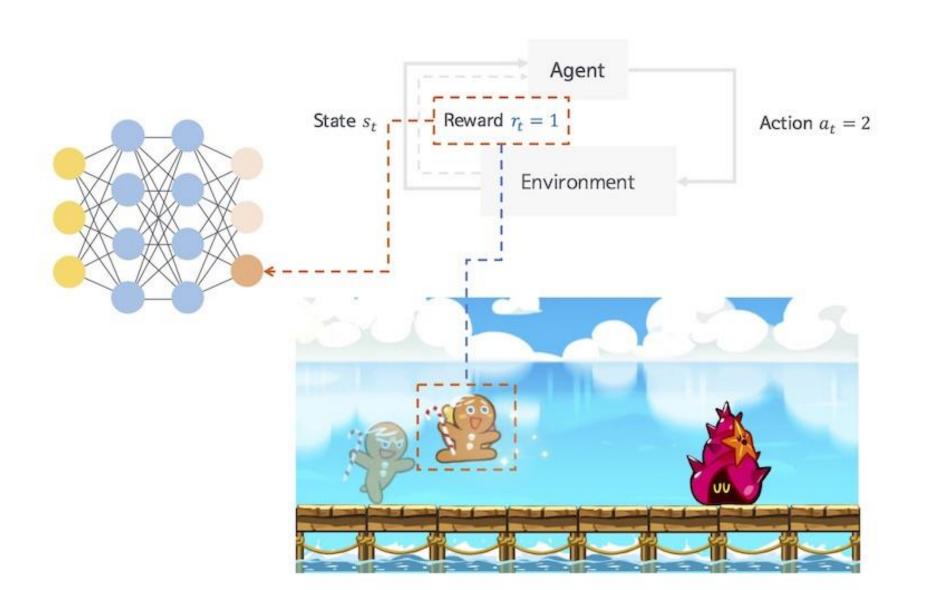
#### 악성코드마다 기대 값이 다를 수 있음

지금까지 구동한 악성코드의 모든 API의 기대 값을 데이터베이스화

어떤 상황에서 어떤 값으로 대응하여 코드커버리지를 크게하는 …

#### 어떤 상황에서 어떤 값으로 대응하여 코드커버리지를 크게하는…

state Action Reward





## 에뮬레이터

## Genetic Algorithm





#### <del>- 운영체제를 에뮬레이션 하고····</del> ≪

<del>- 악성코드가 원하는 API 리턴 값을 잘 전달하며 - </del> ≪

최소한의 API 코딩으로 **악성**코드 행위를 추적할 수 있다면…

**샌드박스**로부터 자유를 외칠 수 있다.





Execution

Space after Filename

Trusted Developer Utilities

Windows Management

Source

Extensions

Launch Agent

Launch Daemon Launchett

LC\_LOAD\_DYLIB Addition

Local Job Scheduling

Group Policy Modification

Hidden Window HISTCONTROL

Indicator Blocking

Hidden Files and Directories

Image File Execution Options

cheduled Task

Service Registry

Setuid and Setgid

Startup Items

Sudo Caching

Sudo

SID-History Injection

### 에뮬레이터 최소한의 API 코딩으로 악성코드 행위를 추적할 수 있다!

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Command And Control	Exfiltration	Impact
11 items	34 items	62 items	32 items	69 items	21 items	23 items	18 items	13 items	22 items	9 items	16 items
Drive-by Compromise	AppleScript	.bash_profile and .bashro	Access Token	Access Token Manipulation	Account Manipulation	Account Discovery	AppleScript	Audio Capture	Commonly Used Port	Automated Exfiltration	Account Access Removal
Exploit Public-Facing	CMSTP	Accessibility Features	Manipulation	Binary Padding	Bash History	Application Window Discovery	Application Deployment	Automated Collection	Communication Through	Data Compressed	Data Destruction
Application	Command-Line Interface	Account Manipulation	Accessibility Features	BITS Jobs	Brute Force	Browser Bookmark Discovery	Software	Clipboard Data	Removable Media	Data Encrypted	Data Encrypted for Impact
External Remote Services	Compiled HTML File	AppCert DLLs	AppCert DLLs	Bypass User Account Control	Credential Dumping	Domain Trust Discovery	Component Object Model and Distributed	Data from Information	Connection Proxy	Data Transfer Size	Defacement
Hardware Additions	Component Object Model	Applnit DLLs	Appinit DLLs	Clear Command History	Credentials from Web	File and Directory Discovery	COM	Repositories	Custom Command and Control Protocol	Limits	Disk Content Wipe
Replication Through	and Distributed COM	Application Shimming	Application Shimming	CMSTP	Browsers	Network Service Scanning	Exploitation of Remote Services	Data from Local System	Custom Cryptographic	Exhitration Over Alternative Protocol	Disk Structure Wipe
Removable Media	Control Panel Items	Authentication Package	Bypass User Account Control	Code Signing	Credentials in Files	Network Share Discovery	Internal Spearphishing	Data from Network	Protocol	Exfiltration Over	Endpoint Denial of Service
Spearphishing Attachment	Dynamic Data Exchange	BITS Jobs	DLL Search Order	Compile After Delivery	Credentials in Registry	Network Sniffing	Logon Scripts	Shared Drive	Data Encoding	Command and Control Channel	Firmware Corruption
Spearphishing Link	Execution through API	Bootkit	Hijacking	Compiled HTML File	Exploitation for Credential Access	Password Policy Discovery	Pass the Hash	Data from Removable Media	Data Obfuscation	Exfiltration Over Other	Inhibit System Recovery
Spearphishing via	Execution through Module Load	Browser Extensions	Dylib Hjacking	Component Firmware	Forced Authentication	Peripheral Device Discovery	Pass the Ticket	Data Staged	Domain Fronting	Network Medium	Network Denial of Service
Service	Exploitation for Client	Change Default File	Elevated Execution with Prompt	Component Object Model	Hooking	Permission Groups Discovery	Remote Desktop	Email Collection	Domain Generation Algorithms	Exfiltration Over Physical Medium	Resource Hijacking
Supply Chain Compromise	Execution	Association	Emond	Historing	Input Capture	Process Discovery	Protocol	Input Capture	Fallback Channels	Scheduled Transfer	Runtime Data Manipulation
Trusted Relationship	Graphical User Interface	Component Firmware	Exploitation for	Connection Proxy	Input Prompt	Query Registry	Remote File Copy	Man in the Browser	Multi-hop Praxy		Service Stop
Valid Accounts	InstaliUtil	Component Object Model Hijacking	Privilege Escalation	Control Panel Items	Kerbereasting	Remote System Discovery	Remote Services	Screen Capture	Multi-Stage Channels		Stored Data Manipulation
	Launcheti	Create Account	Extra Window Memory Injection	DCShadow	Keychain	Security Software Discovery	Replication Through Removable Media	Video Capture	Multiband Communication		System Shutdown/Reboot
	Local Job Scheduling	DLL Search Order	File System	Deobfuscate/Decode Files or Information	LLMNR/NBT-NS	Software Discovery	Shared Webroot		Multilayer Encryption		Transmitted Data
	LSASS Driver	Hijacking	Permissions Weakness	Disabling Security Tools	Poisoning and Relay	System Information Discovery	SSH Hilacking		Port Knocking		Manipulation
	Mshta	Dylb Hjacking	Hooking	DLL Search Order Hijacking	Network Sniffing	System Network Configuration	Taint Shared Content		Remote Access Tools		
	PowerShell	Emond	Image File Execution Options Injection	DLL Side-Loading	Password Filter DLL	Discovery	Third-party Software		Remote File Copy		
	Regsvcs/Regasm	External Remote Services	Launch Daemon	Execution Guardralis	Private Keys	System Network Connections Discovery	Windows Admin Shares		Standard Application		
	Regsvr32	File System Permissions Weakness	New Service	Exploitation for Defense Evasion	Securityd Memory	System Owner/User Discovery	Windows Remote		Layer Protocol		
	Rundli32	NAMES OF THE OWNER			Steal Web Session		Management		Standard Cryptographic		
	Scheduled Task	Hidden Files and Directories	Parent PID Spoofing	Extra Window Memory Injection	Cookie	System Service Discovery			Protocol		
	Scripting	Hooking	Path Interception	File and Directory Permissions Modification	Two-Factor Authentication	System Time Discovery			Standard Non-Application Layer Protocol		
	Service Execution	Hypervisor	Plist Modification	File Deletion	Interception	Virtualization/Sandbox Evasion			Uncommonly Used Port		
	Signed Binary Proxy	Image File Execution	Part Manitors	File System Logical Offsets					Web Service		
	Execution	Options Injection	PowerShell Profile	Gatekeeper Bypass							
	Signed Script Proxy	Kernel Modules and	Process Injection	harman and the state of the sta							





## 감사합니다



최원혁 whchoi@nurilab.com