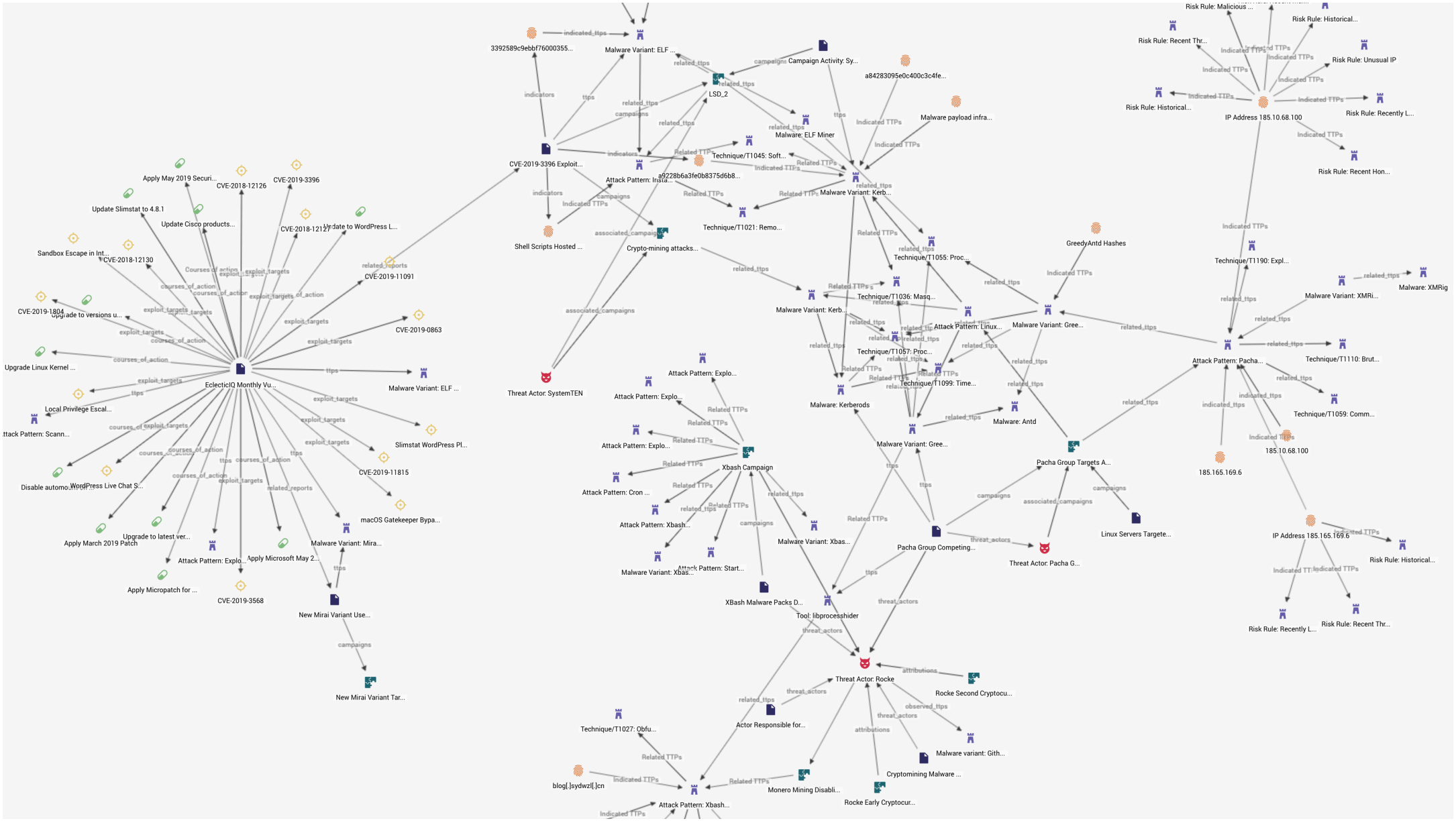
Generating MITRE ATT&CK DNA for groups of actors

Sergey Polzunov,
EclecticIQ

Hypotheses

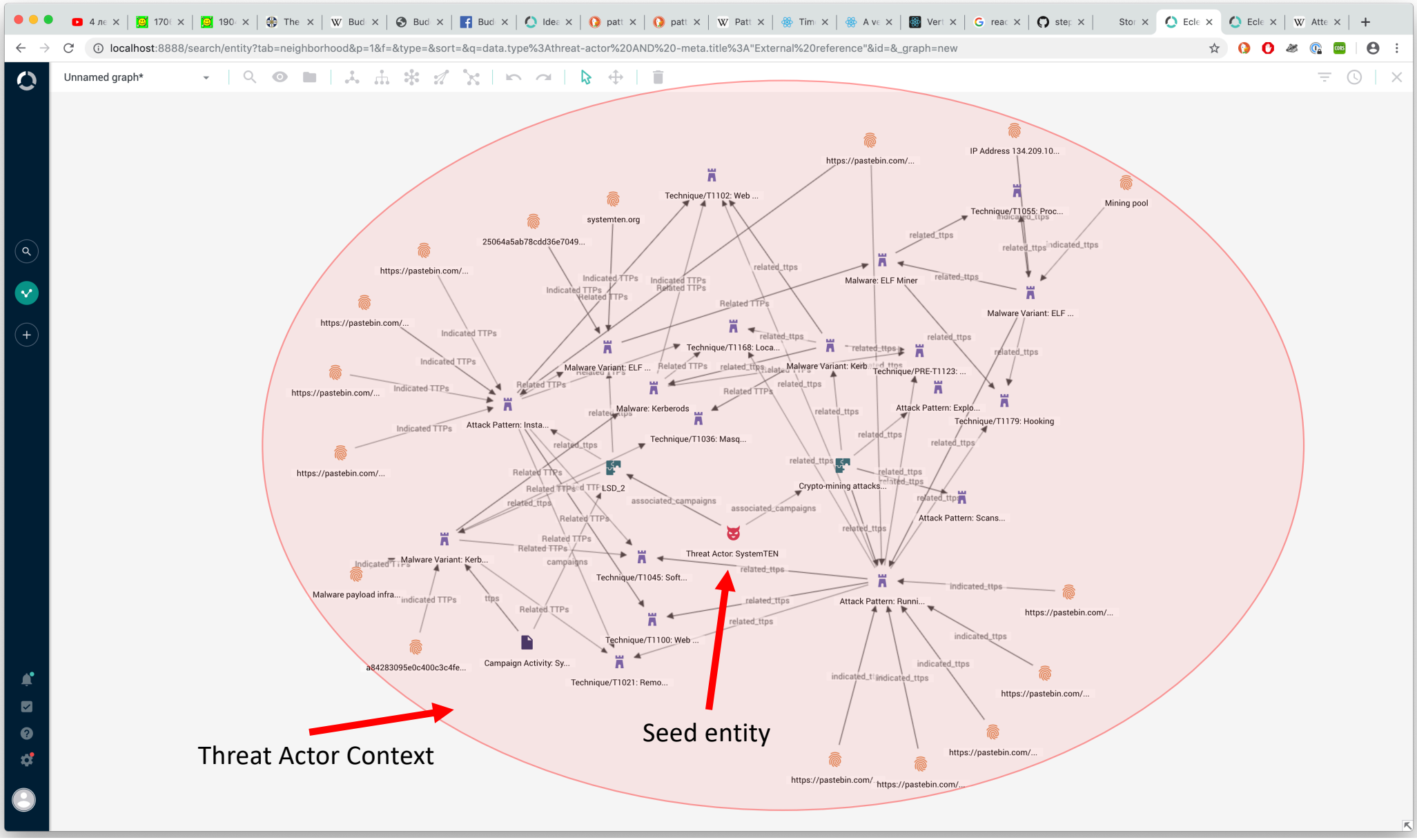
- There are differences in styles of operation on a threat actor / intrusion set level
- These differences can be detected when we compare ATT&CK techniques observed in campaigns
- Subtle differences on a threat actor level become more pronounced when actors are grouped together (by nation state, for example)


EclecticIQ Monthly Vu...
Apply May 2019 Securi...
CVE-2018-12126
CVE-2019-3396
Update Slimstat to 4.8.1
Update Cisco products...
CVE-2018-12127
Update to WordPress L...
Sandbox Escape in Int...
CVE-2018-12130
CVE-2019-11091
CVE-2019-1804
Upgrade to versions u...
CVE-2019-0863
Upgrade Linux Kernel ...
Malware Variant: ELF ...
Local Privilege Escal...
ttack Pattern: Scann...
Slimstat WordPress Pl...
Disable automo...
WordPress Live Chat S...
CVE-2019-11815
macOS Gatekeeper Bypa...
Apply March 2019 Patch
Upgrade to latest ver...
Malware Variant: Mira...
Attack Pattern: Explo...
Apply Microsoft May 2...
Apply Micropatch for ...
CVE-2019-3568
New Mirai Variant Use...
New Mirai Variant Tar...
3392589c9ebbf76000355...
Malware Variant: ELF ...
Campaign Activity: Sy...
a84283095e0c400c3c4fe...
LSD_2
Malware payload infra...
Malware: ELF Miner
CVE-2019-3396 Exploit...
Technique/T1045: Soft...
a9228b6a3fe0b8375d6b8...
Attack Pattern: Insta...
Malware Variant: Kerb...
Technique/T1021: Remo...
Shell Scripts Hosted ...
Crypto-mining attacks...
Threat Actor: SystemTEN
Technique/T1055: Proc...
GreedyAntd Hashes
Technique/T1036: Masq...
Malware Variant: Kerb...
Attack Pattern: Linux...
Malware Variant: Gree...
Technique/T1057: Proc...
Technique/T1099: Time...
Malware: Kerberods
Malware: Antd
Attack Pattern: Explo...
Attack Pattern: Explo...
Attack Pattern: Explo...
Attack Pattern: Explo...
Malware Variant: Gree...
Xbash Campaign
Attack Pattern: Cron ...
Attack Pattern: Xbash...
Malware Variant: Xbas...
Malware Variant: Xbas...
Attack Pattern: Start...
XBash Malware Packs D...
Tool: libprocesshider
Pacha Group Competing...
Pacha Group Targets A...
Threat Actor: Pacha G...
Linux Servers Targete...
Threat Actor: Rocke
Rocke Second Cryptocu...
Actor Responsible for...
Technique/T1027: Obfu...
Malware variant: Gith...
Cryptomining Malware ...
blog[.]sydwzl[.]cn
Monero Mining Disabli...
Rocke Early Cryptocur...
Attack Pattern: Xbash...
Risk Rule: Malicious ...
Risk Rule: Historical...
Risk Rule: Recent Thr...
Risk Rule: Unusual IP
Risk Rule: Historical...
IP Address 185.10.68.100
Risk Rule: Recently L...
Risk Rule: Recent Hon...
Technique/T1190: Expl...
Malware Variant: XMRi...
Malware: XMRig
Attack Pattern: Pacha...
Technique/T1110: Brut...
Technique/T1059: Comm...
185.10.68.100
185.165.169.6
IP Address 185.165.169.6
Risk Rule: Historical...
Risk Rule: Recently L...
Risk Rule: Recent Thr...


EclecticIQ Monthly Vu...
Apply May 2019 Securi...
CVE-2018-12126
CVE-2019-3396
Update Slimstat to 4.8.1
Update Cisco products...
CVE-2018-12127
Update to WordPress L...
Sandbox Escape in Int...
CVE-2018-12130
CVE-2019-11091
CVE-2019-1804
Upgrade to versions u...
CVE-2019-0863
Upgrade Linux Kernel ...
Malware Variant: ELF ...
Local Privilege Escal...
ttack Pattern: Scann...
Slimstat WordPress Pl...
Disable automo...
WordPress Live Chat S...
CVE-2019-11815
macOS Gatekeeper Bypa...
Apply March 2019 Patch
Upgrade to latest ver...
Malware Variant: Mira...
Attack Pattern: Explo...
Apply Microsoft May 2...
Apply Micropatch for ...
CVE-2019-3568
New Mirai Variant Use...
New Mirai Variant Tar...
3392589c9ebbf76000355...
Malware Variant: ELF ...
Campaign Activity: Sy...
a84283095e0c400c3c4fe...
LSD_2
Malware payload infra...
Malware: ELF Miner
CVE-2019-3396 Exploit...
Technique/T1045: Soft...
a9228b6a3fe0b8375d6b8...
Attack Pattern: Insta...
Malware Variant: Kerb...
Technique/T1021: Remo...
Shell Scripts Hosted ...
Crypto-mining attacks...
Technique/T1055: Proc...
GreedyAntd Hashes
Technique/T1036: Masq...
Malware Variant: Kerb...
Attack Pattern: Linux...
Malware Variant: Gree...
Technique/T1057: Proc...
Technique/T1099: Time...
Malware: Kerberods
Malware: Antd
Threat Actor: SystemTEN
Attack Pattern: Explo...
Attack Pattern: Explo...
Attack Pattern: Explo...
Attack Pattern: Explo...
Malware Variant: Gree...
Xbash Campaign
Attack Pattern: Cron ...
Attack Pattern: Xbash...
Malware Variant: Xbas...
Malware Variant: Xbas...
Attack Pattern: Start...
XBash Malware Packs D...
Tool: libprocesshider
Pacha Group Competing...
Pacha Group Targets A...
Linux Servers Targete...
Threat Actor: Pacha G...
Threat Actor: Rocke
Rocke Second Cryptocu...
Actor Responsible for...
Technique/T1027: Obfu...
blog[.]sydwzl[.]cn
Attack Pattern: Xbash...
Monero Mining Disabli...
Rocke Early Cryptocur...
Cryptomining Malware ...
Malware variant: Gith...
Risk Rule: Malicious ...
Risk Rule: Historical...
Risk Rule: Recent Thr...
Risk Rule: Unusual IP
Risk Rule: Historical...
IP Address 185.10.68.100
Risk Rule: Recently L...
Risk Rule: Recent Hon...
Technique/T1190: Expl...
Malware Variant: XMRi...
Malware: XMRig
Attack Pattern: Pacha...
Technique/T1110: Brut...
Technique/T1059: Comm...
185.10.68.100
185.165.169.6
IP Address 185.165.169.6
Risk Rule: Historical...
Risk Rule: Recently L...
Risk Rule: Recent Thr...

Algorithm

- *Requirements:*
  - *Data with connected STIX entities, CTI data model on top of STIX and taxonomies (MITRE ATT&CK, etc)*
- Define seed entities:
  - specific threat actors, grouped in datasets (nation states)
- Generate contexts:
  - Walk the graph from the seeds, identify relevant entities, filter out irrelevant ones, **count taxonomy nodes seen in entities**
- Calculate TF-IDFS for ATT&CK techniques seen in the contexts
- Merge most interesting techniques per actor into a group's techniques set


localhost:8888/search/entity?tab=neighborhood&p=1&f=&type=&sort=&q=data.type%3Athreat-actor%20AND%20-meta.title%3A"External%20reference"&id=&_graph=new
Unnamed graph*
Threat Actor: SystemTEN
Malware: ELF Miner
Malware: Kerberods
Malware Variant: ELF ...
Malware Variant: Kerb...
Malware Variant: Kerb
Malware Variant: ELF ...
Attack Pattern: Insta...
Attack Pattern: Explo...
Attack Pattern: Runni...
Attack Pattern: Scans...
Technique/T1102: Web ...
Technique/T1168: Loca...
Technique/T1036: Masq...
Technique/T1045: Soft...
Technique/T1100: Web ...
Technique/T1021: Remo...
Technique/T1055: Proc...
Technique/PRE-T1123: ...
Technique/T1179: Hooking
Crypto-mining attacks...
LSD_2
Campaign Activity: Sy...
systemten.org
25064a5ab78cdd36e7049...
IP Address 134.209.10...
Mining pool
Malware payload infra...
a84283095e0c400c3c4fe...
https://pastebin.com/...
related_ttps
Related TTPs
Indicated TTPs
indicated_ttps
associated_campaigns
campaigns
Threat Actor Context
Seed entity

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Command And Control	Exfiltration	Impact
Drive-by Compromise	AppleScript	.bash_profile and .bashrc	Access Token Manipulation	Access Token Manipulation	Account Manipulation	Account Discovery	AppleScript	Audio Capture	Commonly Used Port	Automated Exfiltration	Data Destruction
Exploit Public-Facing Application	CMSTP	Accessibility Features	Accessibility Features	Binary Padding	Bash History	Application Window Discovery	Application Deployment Software	Automated Collection	Communication Through Removable Media	Data Compressed	Data Encrypted for Impact
External Remote Services	Command-Line Interface	Account Manipulation	AppCert DLLs	BITS Jobs	Brute Force	Browser Bookmark Discovery	Distributed Component Object Model	Clipboard Data	Connection Proxy	Data Encrypted	Defacement
Hardware Additions	Compiled HTML File	AppCert DLLs	AppInit DLLs	Bypass User Account Control	Credential Dumping	Domain Trust Discovery	Exploitation of Remote Services	Data from Information Repositories	Custom Command and Control Protocol	Data Transfer Size Limits	Disk Content Wipe
Replication Through Removable Media	Control Panel Items	AppInit DLLs	Application Shimming	Clear Command History	Credentials in Files	File and Directory Discovery	Logon Scripts	Data from Local System	Custom Cryptographic Protocol	Exfiltration Over Alternative Protocol	Disk Structure Wipe
Spearphishing Attachment	Dynamic Data Exchange	Application Shimming	Bypass User Account Control	CMSTP	Credentials in Registry	Network Service Scanning	Pass the Hash	Data from Network Shared Drive	Data Encoding	Exfiltration Over Command and Control Channel	Endpoint Denial of Service
Spearphishing Link	Execution through API	Authentication Package	DLL Search Order Hijacking	Code Signing	Exploitation for Credential Access	Network Share Discovery	Pass the Ticket	Data from Removable Media	Data Obfuscation	Exfiltration Over Other Network Medium	Firmware Corruption
Spearphishing via Service	Execution through Module Load	BITS Jobs	Dylib Hijacking	Compile After Delivery	Forced Authentication	Network Sniffing	Remote Desktop Protocol	Data Staged	Domain Fronting	Exfiltration Over Physical Medium	Inhibit System Recovery
Supply Chain Compromise	Exploitation for Client Execution	Bootkit	Exploitation for Privilege Escalation	Compiled HTML File	Hooking	Password Policy Discovery	Remote File Copy	Email Collection	Domain Generation Algorithms	Scheduled Transfer	Network Denial of Service
Trusted Relationship	Graphical User Interface	Browser Extensions	Extra Window Memory Injection	Component Firmware	Input Capture	Peripheral Device Discovery	Remote Services	Input Capture	Fallback Channels		Resource Hijacking
Valid Accounts	InstallUtil	Change Default File Association	File System Permissions Weakness	Component Object Model Hijacking	Input Prompt	Permission Groups Discovery	Replication Through Removable Media	Man in the Browser	Multi-hop Proxy		Runtime Data Manipulation
	Launchctl	Component Firmware	Hooking	Control Panel Items	Kerberoasting	Process Discovery	Shared Webroot	Screen Capture	Multi-Stage Channels		Service Stop
	Local Job Scheduling	Component Object Model Hijacking	Image File Execution Options Injection	DCShadow	Keychain	Query Registry	SSH Hijacking	Video Capture	Multiband Communication		Stored Data Manipulation
	LSASS Driver	Create Account	Launch Daemon	Deobfuscate/Decode Files or Information	LLMNR/NBT-NS Poisoning and Relay	Remote System Discovery	Taint Shared Content		Multilayer Encryption		Transmitted Data Manipulation
	Mshta	DLL Search Order Hijacking	New Service	Disabling Security Tools	Network Sniffing	Security Software Discovery	Third-party Software		Port Knocking		
	PowerShell	Dylib Hijacking	Path Interception	DLL Search Order Hijacking	Password Filter DLL	System Information Discovery	Windows Admin Shares		Remote Access Tools		
	Regsvcs/Regasm	External Remote Services	Plist Modification	DLL Side-Loading	Private Keys	System Network Configuration Discovery	Windows Remote Management		Remote File Copy		
	Regsvr32	File System Permissions Weakness	Port Monitors	Execution Guardrails	Securityd Memory	System Network Connections Discovery			Standard Application Layer Protocol		
	Rundll32	Hidden Files and Directories	Process Injection	Exploitation for Defense Evasion	Two-Factor Authentication Interception	System Owner/User Discovery			Standard Cryptographic Protocol		
	Scheduled Task	Hooking	Scheduled Task	Extra Window Memory Injection		System Service Discovery			Standard Non-Application Layer Protocol		
	Scripting	Hypervisor	Service Registry Permissions Weakness	File Deletion		System Time Discovery			Uncommonly Used Port		
	Service Execution	Image File Execution Options Injection	Setuid and Setgid	File Permissions Modification		Virtualization/Sandbox Evasion			Web Service		
	Signed Binary Proxy Execution	Kernel Modules and Extensions	SID-History Injection	File System Logical Offsets							
	Signed Script Proxy Execution	Launch Agent	Startup Items	Gatekeeper Bypass							
	Source	Launch Daemon	Sudo	Group Policy Modification							
	Space after Filename	Launchctl	Sudo Caching	Hidden Files and Directories							
	Third-party Software	LC_LOAD_DYLIB Addition	Valid Accounts	Hidden Users							
	Trap	Local Job Scheduling	Web Shell	Hidden Window							
	Trusted Developer Utilities	Login Item		HISTCONTROL							
	User Execution	Logon Scripts		Image File Execution Options Injection							
	Windows Management Instrumentation	LSASS Driver		Indicator Blocking							
	Windows Remote Management	Modify Existing Service		Indicator Removal from Tools							
	XSL Script Processing	Netsh Helper DLL		Indicator Removal on Host							
		New Service		Indirect Command Execution							
		Office Application Startup		Install Root Certificate							
		Path Interception		InstallUtil							
		Plist Modification		Launchctl							
		Port Knocking		LC_MAIN Hijacking							
		Port Monitors		Masquerading							
		Rc.common		Modify Registry							
		Re-opened Applications		Mshta							
		Redundant Access		Network Share Connection Removal							
		Registry Run Keys / Startup Folder		NTFS File Attributes							
		Scheduled Task		Obfuscated Files or Information							
		Screensaver		Plist Modification							
		Security Support Provider		Port Knocking							
		Service Registry Permissions Weakness		Process Doppelgänging							
		Setuid and Setgid		Process Hollowing							
		Shortcut Modification		Process Injection							
		SIP and Trust Provider Hijacking		Redundant Access							
		Startup Items		Regsvcs/Regasm							
		System Firmware		Regsvr32							
		Systemd Service		Rootkit							
		Time Providers		Rundll32							
		Trap		Scripting							

Histograms → TF-IDF

- Most popular (for the actor) exotic (across all) techniques used

$$TF\ (term) = Term\ Frequency\ (term) = \frac{\#\ of\ times\ term\ appears\ in\ the\ doc}{total\ \#\ of\ terms\ in\ the\ doc}$$

$$IDF\ (term) = Inverse\ Document\ Frequency\ (term) = log\left(\frac{total\ \#\ of\ doc}{\#\ of\ docs\ containing\ term}\right)$$

$$TF\text{-}IDF\ (term) = TF\ (term) * IDF\ (term)$$

TF-IDFs vectors per actor context

TF-IDF	ID	Technique
2,33289044	1328	Buy domain name
1,57675973	1254	Conduct active scanning
1,57675973	1210	Exploitation of Remote Services
1,26640917	1308	Acquire and/or use 3rd party software services
1,1273593	1438	Alternate Network Mediums
1,05117315	1443	Remotely Install Application
0,94980688	1345	Create custom payloads
0,92665868	1203	Exploitation for Client Execution
0,74150134	1460	Biometric Spoofing
0,74132694	1192	Spearphishing Link
0,7256059	1431	App Delivered via Web Download

Disclaimer

- The data has noise in it:
  - Old techniques use same numeric IDs as new ones
  - Mistyped ATT&CK IDs
  - Incorrectly tagged entities
  - Incorrect relations between entities
- The results here are only from the data we have in our DB
- The methodology and our data is surely full of biases

China

Normalised TF-IDFs	
ID	Tech
1431	App Delivered via Web Download
1476	Deliver Malicious App via Other Means
1433	Access Call Log
1402	App Auto-Start at Device Boot
1262	Enumerate client configurations
1422	System Network Configuration Discovery
1497	Virtualization/Sandbox Evasion
1430	Location Tracking
1254	Conduct active scanning
1210	Exploitation of Remote Services

Simple TF-IDFs	
ID	Tech
1328	Buy domain name
1254	Conduct active scanning
1210	Exploitation of Remote Services
1308	Acquire and/or use 3rd party software services
1438	Alternate Network Mediums
1443	Remotely Install Application
1345	Create custom payloads
1203	Exploitation for Client Execution
1460	Biometric Spoofing
1192	Spearphishing Link

Sublinear TF-IDFs	
ID	Tech
1328	Buy domain name
1203	Exploitation for Client Execution
1193	Spearphishing Attachment
1438	Alternate Network Mediums
1192	Spearphishing Link
1254	Conduct active scanning
1210	Exploitation of Remote Services
1363	Port redirector
1461	Lockscreen Bypass
1308	Acquire and/or use 3rd party software services

China and Russia, techniques per dataset

CHINA	
ID	Tech
1192	Spearphishing Link
1193	Spearphishing Attachment
1203	Exploitation for Client Execution
1210	Exploitation of Remote Services
1254	Conduct active scanning
1262	Enumerate client configurations
1308	Acquire and/or use 3rd party software services
1328	Buy domain name
1345	Create custom payloads
1363	Port redirector
1402	App Auto-Start at Device Boot
1422	System Network Configuration Discovery
1430	Location Tracking
1431	App Delivered via Web Download
1433	Access Call Log
1438	Alternate Network Mediums
1443	Remotely Install Application
1460	Biometric Spoofing
1461	Lockscreen Bypass
1476	Deliver Malicious App via Other Means
1497	Virtualization/Sandbox Evasion

RUSSIA	
ID	Tech
1176	Browser Extensions
1190	Exploit Public-Facing Application
1192	Spearphishing Link
1193	Spearphishing Attachment
1204	User Execution
1219	Remote Access Tools
1328	Buy domain name
1332	Acquire or compromise 3rd party signing certificates
1334	Compromise 3rd party infrastructure to support delivery
1347	Build and configure delivery systems
1350	Discover new exploits and monitor exploit-provider forums
1419	Device Type Discovery
1424	Process Discovery
1430	Location Tracking
1461	Lockscreen Bypass
1465	Rogue Wi-Fi Access Points
1468	Remotely Track Device Without Authorization
1489	Service Stop

What’s next

- Improve the method, compensate for biases
- Automatically highlight potentially interesting techniques to threat analysts conducting analysis
- Similarity estimation
  - Using TF-IDF vectors for calculating similarity between activities of actors
- Use sub-techniques to increase the granularity
- Compare techniques for selected actors with the ones listen in MITRE ATT&CK groups

Questions?

sergey@eclecticiq.com