C-TAS Ecosystem for Cyber Threat Analysis & Sharing in Korea

2017. 12. 7

Sang Wook Seo, KISA/KrCERT
Jung Hee Kim, KISA/KrCERT
Dong Ryun Lee, KISA/KrCERT
Prof. Huy Kang Kim, Korea Univ.
About authors

**Sang Wook Seo (Speaker)**
- General Researcher, National Cyber Intelligence Team, Korea Internet & Security Agency
- Ph.D Course, Graduate School of Information Security, Korea University
- Big Data System & Data Architect, Data Mining & Machine Learning in Security

**Jung Hee Kim**
- Director, Cyber Threat Intelligence Center, Korea Internet & Security Agency
- Director of National & Global Cyber Threat Intelligence Cooperation in Korea

**Dong Ryun Lee**
- Manager, National Cyber Intelligence Team, Korea Internet & Security Agency
- Coordinator of National Cyber Threat Intelligence Network in Korea

**Huy Kang Kim**
- Associate Professor, Graduate School of Information Security, Korea University
- Online Game Security, Fraud Detection System, Network & System Security
Contents

1. C-TAS System
2. C-TEX Structure
3. Big Data in C-TAS
1. C-TAS System
1-1. Introduction to C-TAS System

C-TAS system was developed to prevent the spread of harm from various cyber incidents by collecting, analyzing and disseminating cyber threats.
1-2. Motivation & History

**C-TAS (Cyber Threat Analysis & Sharing) System**
- by KISA (Korea Internet & Security Agency), August 2014

**Motivation**
- NH APT Attack (2011) & 3.20 APT Attack (2013, DarkSeoul)
- Korea Hydro & Nuclear Power Hacking (2014)

**Development**
- 12.05 ~ 12.11 : MMS 1.0 & MML 1.0
- 13.08 ~ 13.12 : MMS 1.1 & MML 1.1
- 13.09 ~ 14.07 : C-TAS 1.0 & C-TAS 1.0
- 15.05 ~ 15.12 : C-TAS 1.1 & C-TEX 1.1 (MMS -> TIMS)
- 16.05 ~ 16.12 : C-TAS 1.2 & C-TEX 1.2 (with STIX 1.2)
- 17.05 ~ 17.12 : C-TAS 2.0 & C-TEX 2.0 (with STIX 2.0)

- C-TAS : Cyber Threat Analysis & Sharing
- C-TEX : Cyber Threat EXpression
- MMS : Malware Management System
- MML : Malware Markup Language
- TIMS : Threat Intelligence Management System
1-3. Collecting Cyber Threat

From KISA & Participants

- **Cyber Threat**: Malware, Malicious Domain/IP, Vulnerability Info and etc
- **Collecting Method**: Agent, Web API, Website

KISA Detection Systems

- Malware
- Domain/IP
- Vulnerability

C-TAS Participants

- Malware
- Domain/IP
- Vulnerability

Cyber Threat Collecting

- Agent
- Web API

C-TAS System

- Collecting Agent
- Web API

Cyber Threat Sharing

- Automatically
- Website
1-4. Disseminating Cyber Threat

To C-TAS Participants

The ways to disseminate cyber threats are:

- Web API to respond to cyber threats in real time
- Website to download & upload cyber threats manually
- STIX/TAXII 2.0 will be supported in 2018

C-TEX & STIX 2.0 (2018)

The ways to disseminate are:

1. Web API (export API) & TAXII (2018)
2. Website (https://cshare.krcert.or.kr)
If you want cyber threats, you must share cyber threats (no free-riding)
You can get the same types of cyber threat you share (type symmetric)
The amount you share decides your grade (4 grades)
Higher grades give you additional information (quality symmetric)
1-6. C-TEX Sample

**C-TEX 1.2 (XML)**

```xml
<mcf>
  <when>
    <date>2015-07-08</date>
    <time>05:36:19</time>
  </when>
  <method>system</method>
  <channel>mcf</channel>
  <source/>
  <comment/>
  <address/>
    <domain>www. .com</domain>
    <ip>211.192.139.</ip>
    <icountry>KR</icountry>
    <type>distribute</type>
    <company/>
    <completed>Y</completed>
    <hosting/>
    <toolkit>CKVIP</toolkit>
  </address>
  <vulnerability>
    <cve>CVE-2013-0422</cve>
    <product>JAVA</product>
  </vulnerability>
  <sample>
    <md5>3da8ef90d78766208088d7fa72</md5>
    <sha256>81ab5d27b5311cc7ee1139a2d11c71e5ae1f974caee0000716c6ae2c25</sha256>
    <ssdeep>1536:Vx17FW/zh0M317T7vJ332gA5FM+AfKdPFY+S1Nf2CqM8XKh0JF4BDu:vlDkdXl3</ssdeep>
    <name>de.exe</name>
    <type>infoleak</type>
  </sample>
</mcf>
```

**C-TEX 2.0 (JSON)**

```json
{
  "when": {
    "date": "2015-07-08",
    "time": "05:36:19"
  },
  "method": "system",
  "channel": "mcf",
  "source": "",
  "comment": "",
  "address": [
    {
      "domain": "www. .com",
      "ip": "211.192.139.",
      "icountry": "KR",
      "type": "distribute",
      "company": "",
      "completed": "Y",
      "hosting": "",
      "toolkit": "CKVIP",
    },
    {
      "vulnerability": [
        {
          "cve": "CVE-2013-0422",
          "product": "JAVA",
        },
      ],
      "sample": {
        "md5": "3da8ef90d78766208088d7fa72e",
        "sha256": "81ab5d27b5311cc7ee1139a2d11c71e5ae1f974caee0000716c6ae2c25",
        "ssdeep": "1536:Vx17FW/zh0M317T7vJ332gA5FM+AfKdPFY+S1Nf2CqM8XKh0JF4BDu:vlDkdXl3",
        "name": "de.exe",
        "type": "infoleak",
      }
    }
  ]
}```
1-7. C-TEX to STIX

C-TEX 1.2 to STIX 1.2

C-TEX 1.2 (XML)

```xml
<mcf>
  <when>
    <date>2015-07-08</date>
    <time>05:36:19</time>
  </when>
  <method>system</method>
  <channel>mcf</channel>
  <source/>
  <comment/>
  <address>
    <domain>www...com</domain>
    <ip>211.192.139.</ip>
    <icountry>KR</icountry>
    <url>http://www...com/dataroom/kk/index.html</url>
  </address>
  <type>distribute</type>
  <company/>
  <completed>Y</completed>
  <hosting>
    <toolkit>CKVIP</toolkit>
  </hosting>
  <vulnerability>
    <cve>CVE-2013-0422</cve>
    <product>JAVA</product>
  </vulnerability>
  <sample>
    <md5>3da8ef9d078766208808d7fa72a</md5>
    <sha256>81ab5d27b5311cc7ee1139a2d11c71e5aecf1974caee0000716cfae2c29</sha256>
    <ssdeep>1536:Vxla7Fw/zh0M3I777zV3J32gKASPM+AfKdPFY+S8FZCqm8khh0JF4BDu:vLDKdXl3p29d</ssdeep>
    <name>de.exe</name>
    <type>infoleak</type>
  </sample>
</mcf>
```

STIX 1.2 (XML)

```xml
<stix:STIX_Package>
  <stix:STIX_Header>
    <stix:Information_Source>
      <stixCommon:Identity
        xsi:schemaLocation="stixCommon:Identity xsd:Identity">
        <stixCommon:Role xref="stivocab:InformationSourceRoleVocab-1.0">Initial Author</stixCommon:Role>
        <stixCommon:Time xref="stivocab:DateTimeVocab-1.0">2015-07-08T00:00:00Z</stixCommon:Time>
      </stixCommon:Identity>
      <stixCommon:Time xref="stivocab:DateTimeVocab-1.0">2015-07-08T00:00:00Z</stixCommon:Time>
    </stix:Information_Source>
    <stixCommon:Type xref="stivocab:STIXTypeVocab-1.0">IP Watchlist</stixCommon:Type>
    < Indicator_Observable>
      <stix:Properties>
        <stixCommon:Observable>
          <stixCommon:Observed_Objects>
            <stix:Properties>
              <stixCommon:Observed_OBJECT>
                <stixCommon:Properties>
                  <stix:Properties>
                    <stixCommon:AddressObject>
                      <stixCommon:Properties>
                        <stix:Properties>
                          <stixCommon:AddressValue>
                            <stix:Properties>
                              <stixCommon:AddressValue>211.192.139.</stixCommon:AddressValue>
                            </stix:Properties>
                          </stixCommon:AddressValue>
                        </stix:Properties>
                      </stixCommon:AddressObject>
                    </stixCommon:Properties>
                  </stixCommon:Properties>
                </stixCommon:Observed_OBJECT>
              </stixCommon:Properties>
            </stixCommon:Observed_Objects>
          </stixCommon:Observable>
        </stix:Properties>
      </stix:Properties>
    </Indicator_Observable>
  </stix:STIX_Header>
</stix:STIX_Package>
```
1-8. Supports for C-TAS Participants

**To Search & Visualize Threats**

**C-TAS Analysis Module: Modified ELK Stack**

C-TAS AM: Tool for C-TAS participants to search and visualize cyber threats easily

**KISA**

C-TAS

- Export API
- Storage

**C-TAS Participant**

1. C-TAS Converter
2. Elasticsearch
3. Kibana

**Users**

1. Logstash is replaced by C-TAS Converter to support C-TEX
2. Elasticsearch helps C-TAS participants to search cyber threats
3. Kibana helps C-TAS participants to visualize cyber threats

---

**KrCERT**
1-8. Supports for C-TAS Participants
1-9. Cyber Threat Use Cases

1. Malicious Domain/IP

For All Participants

C-TAS to Security Solution

- Store the malicious Domain/IPs from C-TAS into Threat DB
- Apply cyber threat information in Threat DB to their security solutions
1-9. Cyber Threat Use Cases

2. Malware Sample

For AV & Security

C-TAS to Antivirus

C-TAS Participants

- Store the malware samples from C-TAS into Threat DB
- Update malware signatures for antivirus using Threat DB
- Detect malware in users’ computer

C-TAS

KISA

Threat DB

Export API

Update Malware Signatures

Antivirus

Malware diagnostics

Users
1-9. Cyber Threat Use Cases

3. Malware Hash

- Store the malware hashes from C-TAS into Threat DB
- Web users upload files to a blog or send files over email
- Compare the file hashes to the malware hashes in Threat DB
2. C-TEX Structure
2-1. Introduction to C-TEX

C-TEX (Cyber Threat EXpression)
- Markup Language to express cyber threats

Motivation
- To make it easy for everybody to share cyber threats
- Even for kids!

Development
- 12.05 ~ 12.11 : MMS 1.0 & MML 1.0
- 13.08 ~ 13.12 : MMS 1.1 & MML 1.1
- 13.09 ~ 14.07 : C-TAS 1.0 & C-TAS 1.0
- 15.05 ~ 15.12 : C-TAS 1.1 & C-TEX 1.1 (MMS -> TIMS)
- 16.05 ~ 16.12 : C-TAS 1.2 & C-TEX 1.2 (with STIX 1.2)
- 17.05 ~ 17.12 : C-TAS 2.0 & C-TEX 2.0 (with STIX 2.0)

C-TAS : Cyber Threat Analysis & Sharing
C-TEX : Cyber Threat EXpression
MMS : Malware Management System
MML : Malware Markup Language
TIMS : Threat Intelligence Management System
2-2. C-TEX Structure

C-TEX (Cyber Threat EXpression)

- Collect Markup Language: Address (Domain/IP), Sample (Malware), Vulnerability (Vulnerability)
- Core Markup Languages: Incident, Domain, Host, Sample, Vulnerability, Adversary

- CML (Collect Markup Language)
  - Address, Sample, Vulnerability
- IML (Incident Markup Language)
  - Details on cyber Incident
- DML (Domain Markup Language)
  - Details on registered Domain
- HML (Host Markup Language)
  - Details on hacked Host
- SML (Sample Markup Language)
  - Details on malware Sample
- VML (Vulnerability Markup Language)
  - Details on Vulnerability info
- AML (Adversary Markup Language)
  - Details on Adversary
2-3. C-TEX Schema
2-4. C-TEXg Structure

AML (Adversary) has relationships with IML (Incident)
IML (Incident) has relationships with HML (Host), SML (Sample), VML (vulnerability)
HML (Host), SML (Malware), VML (Vulnerability) has relationships with each other
HML (Host) has relationship with DML (Domain)
2-5. C-TEXg Schema

Relationships between XMLs
2-6. Internal Sources

From KISA Systems

- Cyber Threat Detection Systems collect cyber threats in CML
- The analysts turn cyber threat information into intelligence in IML, HML, SML, VML, AML

Cyber Threat Detection Systems
- Web Crawler
- DDoS Defense System
- Email Detection System
- Mobile Detection System
- Honeypot/Honeynet
- DNS Sinkhole
- etc.

Threat Intelligence Mngmt. System
- Incident Mngmt. System
- Malware Mngmt. System
- Vulnerability Mngmt. System

KISA Systems

Detection
- Web Crawler
- DDoS Defense
- Email Detection
- Mobile Detection
- Honey/Net/Pot
- DNS Sinkhole

Management
- Intelligence Mngmt.
- Incident Mngmt.
- Malware Mngmt.
- Vulnerability Mngmt.

Indicator
- Address
- Sample
- CML
- Vulnerability
- Message

Profile
- DDoS
- APT
- C2
- DBD
- Phishing
- Pharming
- Keylogging
- Mobile
- Z-Day
- Buffer Overflow
- Anonymous
- ...

C-TAS Participants

Security
Global
E/W Vendor
Finance
Infrastructure
Organization
2-7. C-TEX Use Case (Drive By Download)

Website
- domain3.co.kr/2
  - domain4.co.kr/2
    - wiee.emf
- domain6.org/1
- domain7.co.kr/1
- domain8.co.kr/1
- domain2.or.kr/1

Malware
- domain9.co.kr/1
- domain10.com/1
- domain11.co.kr/1
- vire.emf
- qubn.emf
- qpkj.emf
- vire.emf
- qpkj.emf
- upvd.emf
- upvd.emf
- ookm.emf
- fopo.emf

C2
- 192.187.127.xxx
- domain1.com/1
- domain3.co.kr/1
- domain5.com/1
2-7. C-TEX Use Case (Drive By Download)

Same Hosting Company
- domain5.co.kr/1
- domain6.co.kr/1
- domain7.com/1

Same Domain
- domain2.or.kr/1
- domain2.or.kr/2
- domain2.or.kr/3

Website
- domain1.com/1
- domain3.co.kr/1
- domain4.com/1
- domain8.co.kr/1
- domain9.co.kr/1
- domain10.com/1
- d11.co.kr/1
- d12.co.kr/1
- d13.co.kr/1
- domain14.org/1
- domain15.or.kr/1

Malware
- upvd.emf
- ookm.emf
- fopo.emf
- eyip.exe
- hlkk.exe
- asqw.emf
- qwas.emf
- kasm.exe
- srab.emf

C2
- 192.187.127.xxx
- 121.115.165.xxx
2-8. C-TEXg Use Case (Drive By Download)
3. Big Data in C-TAS
3-1. Big Data Platform in C-TAS

C-TAS System Architecture

Contents Management System (Ajax)

Drawing Tool (HTML5)

Graph Visualization (JavaScript)

Web Application Server (Spring Framework)

RBAC (RDBMS)

OLAP (RDBMS)

Document DBMS (mongoDB)

Search Engine (Elasticsearch)

Graph OLAP (Spark GraphX)

Graph DBMS (Neo4j)

Application

Search Engine

DFS (HDFS)

Log Files

Standardization

ETL (Spark Scala)

Sources

Disseminating
library(sna)
edgelist <- read.csv(file="edgelist.csv", header=TRUE, sep=",")
nodelist <- read.csv(file="nodelist.csv", header=TRUE, sep=",")
edgelist <- as.matrix(edgelist)
nodelist <- as.matrix(nodelist)
adacency <- matrix(data=0, nrow=25, ncol=25)
rownames(adjacency) <- nodelist
colnames(adjacency) <- nodelist
adjacency[edgelist] <- 1
centrality <- degree(data=adjacency, gmode="digraph", diag=FALSE, mode="freeman", recscale=FALSE)

plot(data = plot_data, lab = "node", lab.cex = 0.8, edge.col = "grey", displaylabels = TRUE, vertex.col = "white", centrality = centrality, main="KRCERT")

library(TNA)
edgelist <- read.csv(file="edgelist.csv", header=TRUE, sep=",")
nodelist <- read.csv(file="nodelist.csv", header=TRUE, sep=",")
edgelist <- as.matrix(edgelist)
nodelist <- as.matrix(nodelist)
adacency <- adjacency[edgelist]
total_weights <- sum(adjacency)

centrality <- degree(data=adjacency, gmode="digraph", diag=FALSE, mode="freeman", recscale=FALSE)

plot(adjacency, label.cex=0.8, edge.col="grey", displaylabels=TRUE, vertex.col="white", lab="node", lab.cex=0.8, main="KRCERT"

plot_data <- data.frame(nodelist, centrality)
plot_data <- plot_data[order(centrality),]

barplot(plot_data[,2], names.arg=plot_data[,1], col=ifelse(plot_data[,2]<3, "red", "blue"), xlab="node", ylab="centrality", main="KRCERT")
Thank you!