Master 2 SeCReTS

TP Extraction et Analyse de malware

Première partie

Contexte

1 Objectif du TP

Durant ce TP, nous allons analyser un enregistrement (appelé aussi trace) réseau d'une machine aux alentours de sa compromission.

Après une analyse globale de la trace, nous mettrons en évidence deux compromissions potentielles, qui seront alors analysées en profondeur.

ATTENTION ! Les binaires présents dans ce TP sont de vrais samples. Ne les exécutez SURTOUT PAS en environnement non contrôlé (ie. en dehors d'une machine virtuelle), et de préférence sous Linux.

2 Analyse globale

1. Récupérez le fichier de trace réseau à l'adresse `http://moutane.net/uvsq/UE306-Forensics/TP2/trace.cap.zip`, mot de passe « infected ». Avec quoi le visualisez-vous ?
2. Identifiez la machine ciblée et donnez quelques détails, comme son système d'exploitation (vraisemblablement).
3. Quels types de traffic distinguez-vous ?
4. Pour la partie HTTP, donnez la liste des domaines contactés.
5. Un domaine revient plus souvent que les autres.
   (a) Lequel ?
   (b) À quoi sert-il ?
6. Réalisez les mêmes manipulations avec Tshark.
7. (Bonus) Réalisez les mêmes manipulations avec Scapy.
8. Avec Scapy ou Tshark, ecrivez un script qui extrait les adresses accédées via le « proxy ».

Deuxième partie

E-mail

3 Extraction

Dans cette partie, vous aurez besoin d'extraire les pages Web et les documents transmis par le protocole HTTP. Le plus simple est d'utiliser la commande `tcpflow`, par exemple :

```
$ tcpflow -r trace.cap -e http -o <dossier_de_sortie> 'host <adresse_ip>
```

1. Durant cette session, un webmail est utilisé. Lequel ?
2. Quel est le compte utilisé ? Donnez le nom du compte et son mot de passe associé. *(Recherchez le contenu des requêtes POST.)*
3. Un échange d'e-mail est effectué. Résumez-le, en fournissant le contenu des e-mails.
4. L'e-mail lu contient une pièce jointe.
   (a) A-t-elle été téléchargée ?
   (b) Vous semble-t-elle légitime ?

4 Analyse de la pièce jointe

Dans cette partie, nous allons analyser la pièce jointe pour statuer sur la tentative de compromission, voir sur la compromission effective.

1. Récupérez le contenu de la pièce jointe.
2. Quel est le type de fichier ? *(Utilisez la commande file.)*
3. Quelles astuces sont utilisées pour tromper l'utilisateur ?
4. Le binaire est-il packé ? Justifiez votre réponse. Dans le cas où il serait packé, pouvez-vous le dépacker ? Si oui, faites-le.
5. Identifiez le payload final. Vous pouvez, par exemple, vous baser sur les chaînes de caractères contenues dans le fichier. À quelle catégorie appartient-t-il ?
6. Cette famille de payload étant très connue, des études ont été faites dessus. Trouvez le début du protocole de discussion avec le C & C. Pouvez-vous en déduire si le binaire a été lancé ?
7. Identifiez le C & C. Que remarquez-vous ?
8. (Bonus) Des vulnérabilités ont été trouvées pour cette famille et ont été publiées.
   (a) Décrivez-en une permettant d'accéder à un fichier sur le C & C.
   (b) La vulnérabilité semble-t-elle présente dans notre cas d'étude ?
   (c) À quoi pourrait-elle servir ?

Troisième partie

Cyber-criminalité

5 Forensic réseau

1. Durant cette session réseau, le site `perdu.com` est consulté. Selon vous, est-ce la première fois que ce site est accédé ?
2. Décrivez l'enchaînement d'évènements qui en découle. Ce schéma est-il connu ? Nommez les différentes étapes.
3. Récupérez les étapes mentionnées dans la dernière question.
   (a) Désobsurcissez les premières étapes (HTML/JS) et analysez leurs contenus.
   (b) Cela confirme-t-il le schéma trouvé plus haut ?
4. Récupérez le shellcode obtenu. D'après le schéma identifié, l'exploit a-t-il fonctionné ? Identifiez la vulnérabilité utilisée.

6 Reverse Engineering

L'information suivante vous est donnée : au lancement du shellcode, le registre EAX pointe sur le point d'entrée du shellcode.

1. Le shellcode obtenu dans la section précédente a une particularité.
   (a) Laquelle ?
   (b) Comment la justifier ?
   (c) Quelle est la conséquence sur son contenu ?
2. D'après la question précédente, une étape préliminaire à l'analyse statique complète est nécessaire.
   (a) Réalisez cette étape. Quelle approche utilisez-vous ?
   (b) Cette étape étant faite, on voit apparaître la configuration du shellcode. Que pouvez-vous en déduire sur son fonctionnement ?
   (c) Peut-on s'arrêter là ?
3. Analysez le shellcode et décrivez son fonctionnement. À quoi sert-il ?
4. Validez ou infirmez la réponse à la dernière question de la section 5.

Les outils suivant pourront vous servir :
- OllyDBG : Débugger Windows
- Elfesteem : Bibliothèque de reconstruction de conteneur PE
- `http://jsbeautifier.org/` : Indentation automatique de code Javascript
- FireBug ou la console développeur Firefox : Exécution interactive de Javascript
- SpiderMonkey : Engine Javascript, notamment pour Python
- Miasm : Instrumentation et exécution sandboxée de binaire