শিখবো ম্যালওয়্যার এনালাইসিস

রাহাদুল ইসলাম রাহাদ

শুরুতেই বলে রাখছি, আমাদের যাত্রাটি মোটেও সহজ হবে না। কারণ যাত্রাপথে অনেক বাধা আসবে। সেসকল বাধা বিপত্তি দেখে অনেকেই মাঝপথে হাল ছেড়ে দিবে। তাই আগেই বলছি ভালো ভাবে চিন্তা করে আগামীতে পথ চলার জন্য আপনার লক্ষ্য স্থির করুন।

লক্ষ্য স্থির করার পর আপনার দ্বিতীয় কাজ আপনার ভিত্তি (বেসিক'স) আরো মজবুত করা। ম্যালওয়্যার এনালাইজিং সম্পর্কে আপনার বেসিক নলেজটুকু থাকা আবশ্যক যাতে করে আপনার আগামী পথ চলায় বড় সর বাধার সম্মুক্ষিন হতে না হয়। আমরা কিন্তু হুট করেই কোনো একটি ম্যালওয়্যার নিয়ে সেটিকে এনালাইসিস করা শুরু করতে পারবো না। আমাদেরকে প্রোগ্রামিং থেকে শুরু করে নেটওয়ার্কিং পর্যন্ত অনেক কিছু শিখতে হবে। সো চলুন শুরু করা যাক আমাদের যাত্রা.....!

প্রথম ধাপ – প্রোগ্রামিং শিখাঃ

সি প্রোগ্রামিংঃ

আপনাকে যাত্রা শুরু করতে হবে সি প্রোগ্রামিং দিয়ে। কারণ আপনি যে ম্যালওয়্যারগুলো এনালাইসিস করতে চাচ্ছেন তারা তাদের সোর্স কোডের মাধ্যমে পয়েন্টারের ধারনাটি বেশি ব্যবহার করে আর আপনি একবার যখন পয়েন্টার এর সাথে ফেমিলিয়ার হয়ে যাবেন তখন ম্যালওয়্যার এনালাইস্ট হিসেবে জব করা আপনার জন্য আরো সহজ হয়ে যাবে। ম্যালওয়্যার সি প্রোগ্রামিং এর অনেক লাইব্রেরি ইউজ করে থাকে কারণ বর্তমানে এখনো অনেক অপারেটিং সিস্টেম সি প্রোগ্রাম এর উপর দাঁড়িয়ে আছে। অতএব অপারেটিং সিস্টেম এর এপিআই ডকুমেন্টেশনসমুহ অনেকাংশেই সি কেন্দ্রিক। আপনাকে সি প্রোগ্রাম এর মৌলিক বিষয়বস্তুগুলো ভালোভাবে আয়ত্তে আনতে হবে। তবে প্রফেশনাল হওয়ার প্রয়োজন নেই। সি প্রোগ্রামিং শিখার জন্য আপনি আনিসুল ইসলামের সি প্রোগ্রামিং প্লেলিস্টটি দেখতে পারেন।

https://www.youtube.com/watch?v=6nOavbvFvbY&list=PLgH5QX0i9K3pCMBZcul1fta6UivHDbXvz

পাইথন প্রোগ্রামিংঃ

এনালাইজিং এর ক্ষেত্রে আপনাকে খুব স্বল্প সময়ের মধ্যে সাধারণ কাজগুলোকে অটোমেট করতে হবে এবং আপনি যেভাবে চান সেভাবে ম্যালওয়্যারকে ম্যানুপুলেট করতে হবে । আর এসব কাজের জন্য স্বল্প লাইনে কোডিং করে টুলস বানানোর মতো প্রোগ্রামিং ভাষা প্রয়োজন। পাইথন ঠিক এমনই একটি প্রোগ্রামিং ভাষা। তাই এটির কোর নলেজ আপনার থাকতে হবে। পাইথন প্রোগ্রামিং শিখার জন্য আপনি আনিসুল ইসলামের পাইথন প্রোগ্রামিং প্লেলিস্টটি দেখতে পারেন।

https://www.youtube.com/watch?v=xjcCi6Tzfxw&list=PLgH5QX0i9K3rz5XqMsTk41_j15_6682BN

এসেম্বলি ল্যাংগুয়েজঃ

সি এবং পাইথন প্রোগ্রামিং শিখা শেষে এবার আপনাকে এসেম্বলি ল্যাংগুয়েজ শিখতে হবে। কারণ যখন আপনি একটি ম্যালওয়্যার এর রিভার্স ইঞ্জিনিয়ারিং করবেন তখন আপনার বেশিরভাগ সময়ই ব্যয় হবে এসেম্বলি ল্যাংগুয়েজ এর সাথে। সো এসেম্বলি ল্যাংগুয়েজ এর পারদর্শিতা আপনার সাক্সেস রেটকে আরো অনেকগুন বাড়িয়ে তুলতে সহায়তা করবে। এসেম্বলি ল্যাংগুয়েজ শিখার জন্য আপনি নিচের প্লেলিস্টটি দেখতে পারেন।

https://www.youtube.com/watch?v=W9G6JgrQZ5U&list=PL8mraTOYjX3yNe0h3NwvFLgObiG0QLXRQ

সি++ প্রোগ্রামিংঃ

প্রথমে যদিও আমরা সি প্রোগ্রামিং শিখেছি, কিন্তু দুঃখের বিষয় হলো যে সি (**C**) অব্জেক্ট অরিয়েন্টেড প্রোগ্রামিং (**OOP**) না। প্রায় সব ম্যালওয়্যারই সি এবং সি++ প্রোগ্রাম দিয়ে তৈরি করা হয়ে থাকে। আপনি বলতে পারেন যে জাভা দিয়েও তো তৈরি করা যায়। হুম, তৈরি করা যায়। তবে জাভা দিয়ে তৈরি ম্যালওয়্যারকে খুব সহজে রিভার্স করা সম্ভব। তাই সি এবং সি++ এর কোনো বিকল্প নেই। সি++ প্রোগ্রামিং শিখতে নিচের প্লেলিস্টটি দেখতে পারেনঃ

https://www.youtube.com/watch?v=0T4mPpbNs_8&list=PLgH5QX0i9K3q0ZKeXtF--CZ0PdH1sSbYL

Win32 Programming:

যখন আপনি উইন্ডোজ কার্নেল শিখতে যাবেন তখন আপনার win32 programming সম্পর্কে ভালো জ্ঞান থাকা লাগবে। যেমনঃ কিভাবে উইন্ডোজ এপিআই ইউজ করা হয়ে থাকে ইত্যাদি।

Win32 Programming শিখতে নিচের প্লেলিস্টটি দেখতে পারেনঃ

https://www.youtube.com/watch?v=8GCvZs55mEM&list=PLWzp0Bbyy_3i750dsUj7yq4JrPOIUR_NK

দ্বিতীয় ধাপ – কম্পিউটার ও অপারেটিং সিস্টেমঃ

কম্পিউটার আর্কিটেকচারঃ

আপনাকে জানতে হবে কিভাবে প্রসেসর আপনার কম্পিউটার সিস্টেমে কাজ করে। কারণ প্রসেসরই নির্ধারণ করে থাকে যে আপনার অপারেটিং সিস্টেম কিভাবে আপনার কম্পিউটার সিস্টেমকে অপারেট করবে। এক্ষেত্রে আপনি নিচের প্লেলিস্টটি ফলো করতে পারেনঃ

https://www.youtube.com/watch?v=Ol8D69VKX2k&list=PLBlnK6fEyqRgLLlzdgiTUKULKJPYc0A4q

অপারেটিং সিস্টেমঃ

অপারেটিং সিস্টেম কম্পিউটার সিস্টেম এর সাথে কমিউনিকেশন এর জন্য যে ফাংশন ইউজ করে থাকে ম্যালওয়্যার ঠিক সেই ফাংশনটিকে টার্গেট করে থাকে। সো কিভাবে অপারেটিং সিস্টেম তার কাজ সম্পন্ন করে থাকে এই বিষয়ে খুব ভালো থিওরিটিক্যাল জ্ঞান থাকতে হবে। অপারেটিং সিস্টেম সম্পর্কে জানতে নিচের প্লেলিস্টটি দেখতে পারেনঃ

https://www.youtube.com/watch?v=vBURTt97EkA&list=PLBlnK6fEyqRiVhbXDGLXDk_OQAeuVcp2O

নেটওয়ার্কিংঃ

Basic Networking -

যেহেতু ম্যালওয়্যার কোনো না কোনো ভাবে ইন্টারনেট এর মাধ্যমে কমিউনিকেট করে থাকে তাই আমাদের নেটওয়ার্কিং সম্পর্কে ভালো জ্ঞান থাকা প্রয়োজন। কারণ বেশিরভাগ ম্যালওয়্যার-ই কিন্তু ইন্টারনেট এর মাধ্যমে ছড়ায়। সো নেটওয়ার্কিং জানা থাকলে আপনি সেই ম্যালওয়্যার এর

ফুটপ্রিন্ট খুব সহজে বের করতে সক্ষম হবে। অতএব, আপনার স্কিলসগুলোর লিস্টে কম্পিউটার নেটওয়ার্কিং এর জ্ঞান থাকাটাও অপরিহার্য। নেটওয়ার্কিং শিখার জন্য নিচের প্লেলিস্টটি দেখতে পারেনঃ

https://www.youtube.com/watch?v=JFF2vJaN0Cw&list=PLxCzCOWd7aiGFBD2-2joCpWOLUrDLvVV_

Advance Networking -

আপনি এখন নেটওয়ার্কিং সম্পর্কে বেসিকটুকু জানেন। এবার পরের ধাপ হলো ম্যালওয়্যার এনালাইসিসের জন্য নেটওয়ার্কিংকে ব্যাবহার করা। তার জন্য প্র্যাক্টিক্যাল প্যাকেট এনালাইসিস শিখতে হবে। প্যাকেট এনালাইসিস শিখার জন্য নিচের প্লেলিস্টটি দেখতে পারেনঃ

https://www.youtube.com/watch?v=TkCSr30UojM&list=PLu02DfizZn08hmDRngo3_SdHoOqK5PHrB

তৃতীয় ধাপঃ

কম্পাইলারঃ

যখন আপনি একটি ম্যালওয়্যার এর রিভার্স ইঞ্জিনিয়ারিং করবেন তখন শেষ পর্যায়ে গিয়ে হয়তো কোডের এমন কিছু অংশ চোখে পড়বে যা আপনার বোধগম্য হচ্ছে না। কিন্তু আপনি যদি জানেন কিভাবে কম্পাইলার কাজ করে এবং কিভাবে তারা আপনার কোডকে মেশিন ল্যাংগুয়েজ এ অনুবাদ করে তাহলে আপনি খুব সহজেই বুঝতে পারবেন আপনার বোধগম্য না হওয়া কোডটির কাজ কি। প্র্যাক্টিক্যাল রিভার্স ইঞ্জিনিয়ারিং এর জন্য নিচের প্লেলিস্টটি দেখতে পারেনঃ

https://www.youtube.com/watch?v=Nv-GTg3uICE&list=PL-DxAN1jsRa9151ezNuCbh7UkGS0bMPdw

Obfuscation:

একজন ম্যালওয়্যার ডেভেলপার কখনো চাইবে না যে তার উদ্দেশ্য হাসিল হওয়ার আগেই তার ম্যালওয়্যারটি কেউ এনালাইসিস করে তার জন্য এন্টি-ম্যালওয়্যার তৈরি করে ফেলুক। আর তাই তারা তাদের ম্যালওয়্যার এর কোডে Obfuscation ব্যবহার করে থাকে যাতে করে ম্যালওয়্যার এনালাইসিস করাটা আরো কঠিন হয়ে যেতে পারে। তাই একজন ম্যালওয়্যার এনালাইস্ট হিসেবে আপনার এধরনের বিশেষ কিছু বিষয়ে জ্ঞান থাকা আবশ্যক। যেমনঃ Obfusctaion, Watermarking and Tamperproofing etc. এক্ষত্রে আপনি নিচের বইটি পড়ে দেখতে পারেনঃ

https://www.amazon.com/Surreptitious-Software-Obfuscation-Watermarking-Tamperproofing/dp/0321549252

-PE (Portable Executable):

উইন্ডোজ অপারেটিং সিস্টেম কোনো প্রোগ্রামকে রান করতে Portable Executable (PE) ফাইল ফরমেট ব্যবহার করে থাকে। PE ফাইল ফরমেট মূলত একটি ডাটা স্ট্রাকচার যাতে উইন্ডোজ অপারেটিং সিস্টেম লোডারের জন্য আবৃত এক্সিকিউটেবল কোড পরিচালনার জন্য প্রয়োজনীয় তথ্য থাকে। সো একজন ম্যালওয়্যার এনালাইস্ট হিসেবে আপনাকে এই ফাইল ফরমেট এর সাথে পরিচিত থাকতে হবে ।

বিস্তারিত জানতে নিচের পোস্টটি পরুনঃ

https://resources.infosecinstitute.com/topic/2-malware-researchers-handbook-demystifying-pe-file/

Kernel:

একটি প্রোগ্রাম যা আপনার কম্পিউটার সিস্টেমের সফটওয়্যার এবং হার্ডওয়্যার এর মধ্যে যোগাযোগ ব্যবস্থাকে নিয়ন্ত্রন করে। কিছু কিছু ম্যালওয়্যার নিজেকে আরো বেশি গোপন এবং পারসিস্টেন্ট (**Persistent**) থাকার জন্য এই কার্নেল এর সুবিধা নিয়ে থাকে। যদি আপনি ম্যালওয়্যার এনালাইস্ট হিসেবে আপনার ক্যারিয়ারে অগ্রসর হতে চান তবে কার্নেল সম্পর্কে জ্ঞান থাকা আবশ্যক। উইন্ডোজ কার্নেল প্রোগ্রামিং শিখতে নিচের প্লেলিস্টটি দেখতে পারেনঃ

https://www.youtube.com/watch?v=XUlbYRFFYf0&list=PLZ4EgN7ZCzJx2DRXTRUXRrB2njWnx1kA2

অন্তিম পর্বঃ

প্রোগ্রামিং থেকে শুরু করে কার্নেল পর্যন্ত অনেক কিছুই শিখা হলো। এবার আমরা সরাসরি ম্যালওয়্যার এনালাইসিস শিখার জন্য প্রস্তুত। এই পর্বে কিছু বই এবং ব্লগ থেকে আমরা শিখবো। এবং ম্যালওয়্যার এনালাইসিস করার জন্য ল্যাব সেটআপ করা থেকে শুরু করে বিভিন্ন রিসোর্স দেখিয়ে দিবো। শুরু করা যাক...!

1. **Practical Malware Analysis** বইটি নিজের সংগ্রহে রাখুন। **Practical Malware Analysis** বইটি ম্যালওয়্যার এনালাইসিস শিখার জন্য সবচেয়ে বেস্ট বই। কারণ এটি আপনাকে ম্যালওয়্যার এনালাইসিস এর একদম (Core) থেকে শিখাবে এবং আপনাকে অনেক ধরনের জেনারেল টেকনিক'স এর সাথে পরিচয় করিয়ে দিবে ও ম্যালওয়্যার এনালাইসিস এর ক্ষেত্রে যেসকল টুলস ব্যবহার করার প্রয়োজন হয় তাদের ব্যবহার শিখাবে।

2. কিছু ম্যালওয়্যার এর স্যাম্পল (Samples) ডাউনলোড করুন এবং তাদের এনালাইজ করার চেষ্টা করুন। ম্যালওয়্যার এর স্যাম্পলস ডাউনলোড করতে **Github Repo** টি ফলো করতে পারেন।

3. এবার যেই বইটি আপনার পড়া প্রয়োজন তা হলো **Malware Analyst's Cookbook** । এই বইটি ম্যালওয়্যারকে বুঝার জন্য এবং ম্যালওয়্যার এনালাইসিস এর জন্য টুলস তৈরিতে অনেক বেশি সাহায্য করবে।

4. **Reverse Engineering (RE):**

   আগেই বলেছিলাম যে আপনার বেশিরভাগ সময়ই কিন্তু ব্যয় হবে বাইনারি এনালাইজিং এবং এসেম্বলি ল্যাংগুয়েজ এর পেছনে । তাই আপনার অবশ্যই রিভার্স ইঞ্জিনিয়ারিং স্কিলস থাকতে হবে। **begin.re** সাইটটি রিভার্স ইঞ্জিনিয়ারিং শিখা শুরু করার জন্য ভালো প্লাটফর্ম।

   Reversing: Secrets Of Reverse Engineering বইটিও রিভার্স ইঞ্জিনিয়ারিং শিখতে এবং জানতে অনেক সাহায্য করবে

Tools And Environment:

ম্যালওয়্যার অবশ্যই ক্ষতিকর এটি ভুলে গেলে চলবে না। তাই নিজের কম্পিউটার সিস্টেম এর যাতে কোনো ক্ষতি না হয় তার খেয়াল রাখতে হবে। এজন্য আপনাকে ম্যালওয়্যার এনালাইসিস করার প্রি-ইনস্টল্ড সকল টুলস সহ একটি ভার্চুয়াল মেশিন তৈরি করে নিতে হবে, যেখানে আপনি ম্যালওয়্যার এর স্যাম্পলগুলোকে রেখে নিশ্চিন্তে এনালাইজ করতে পারেন।

বিস্তারিতঃ

- Malware analysis for N00bs – part 1: malware and the tools for its analysis (slides)
- Malware Analysis Virtual Machine – by OALabs
- Creating a Simple Free Malware Analysis Environment – by MalwareTech
- Reviews of various tools for reverse engineering

Learning Tools:

আপনাকে IDA, Ghidra, BinaryNinja, Olly Dbg, ImmunityDbg, x64dbg ইত্যাদি টুলস এর ব্যবহার শিখতে হবে। এসব টুলস অনেক ইউজফুল এবং অনেক এডভান্স। নিচে কিছু কোর্স এর লিংক দেয়া হলো যেখানে আপনি টুলসগুলোর সাথে পরিচিত হতে পারবেন।

- TiGa's course on IDA Pro
- Introduction to WinDbg by Anand George

Some More Malware Samples:

- https://malwarebreakdown.com/
- https://www.malware-traffic-analysis.net/
- https://virusshare.com/
- http://contagiodump.blogspot.com/
- http://thezoo.morirt.com/
- https://github.com/ytisf/theZoo
- https://malshare.com
- http://dasmalwerk.eu/
- http://www.virusign.com/
- https://zeltser.com/malware-sample-sources/

Malware Trackers:

নিচে কিছু ম্যালওয়্যার ট্র্যাকার এর লিংক দেয়া হলো যেখানে আপনি লেটেস্ট ম্যালওয়্যার এর লাইভ লিংক এবং আরো অনেক তথ্য পাবেন।

- https://tracker.fumik0.com
- http://benkow.cc
- http://vxvault.net/
- http://cybercrime-tracker.net/

Exercises:

রিভার্সিং হলো একটি আর্ট যা আপনি শুধু মাত্র প্র্যাক্টিস করে করেই শিখতে পারবেন। তাই আমি আপনাকে সরাসরি প্র্যাক্টিস করার পরামর্শ দিবো। নিচে কিছু রাইট-আপ'স (Write-ups) দেয়া হলো যা ফলো করে আপনারা প্র্যাক্টিস করতে পারবেন।

- Beginner Malware Reversing Challenges (by Malware Tech)
- Malwarebytes CrackMe #1 + tutorial
- Malwarebytes CrackMe #2 + list of write-ups
- Malwarebytes Crackme #3 + list of write-ups
- https://crackmes.one/ – various crackmes to help you exercise reversing
- "Nightmare" – a reverse engineering course created around CTF tasks

Malware Unpacking:

ম্যালওয়্যার ডেভেলপাররা এন্টি-ভাইরাসকে ধোকা দিতে বিভিন্ন ধরনের কৌশল অবলম্বন করে থাকে। আর এই কৌশলগুলোর একটি হলো ম্যালওয়্যার প্যাকিং। ম্যালওয়্যার সাধারণত প্যাকড (Packed) অবস্থায় থাকে এবং ম্যালওয়্যারটির কোর থেকে এনালাইজ করার জন্য এটিকে আনপ্যাক (Unpack) করে নিতে হয়। ম্যালওয়্যার আনপ্যাকিং হলো ম্যালওয়্যার এনালাইজ শুরু করার প্রথম ধাপ। যতক্ষন না আপনি একটি ম্যালওয়্যারকে আনপ্যাক করবেন ততোক্ষন আপনি প্যাক করা ম্যালওয়্যারটির মৌলিক কার্যাবলি যেমন স্ট্রিংস (Strings), এপিআই (Api) এর ব্যবহার, ফাইলের ইম্পোর্ট এক্সপোর্ট (Import/ Export) ইত্যাদি বিষয় ইনুমিরেট (Enumerate) করতে পারবো না। তাই আমাদের ম্যালওয়্যার আনপ্যাকিং শিখতে হবে। ম্যালওয়্যার আনপ্যাকিং শিখতে নিচের প্লেলিস্টটি ফলো করতে পারেনঃ

https://www.youtube.com/watch?v=KvOpNznu_3w&list=PL3CZ2aaB7m83eYTAVV2knNgIB8I4y5QmH

Virtualization-Based Obfuscation:

ম্যালওয়্যার ডেভেলপাররা প্রায়ই তাদের কোড এর ক্ষতিকারক কার্যকারিতাকে লুকানোর জন্য Code Obfuscation ব্যবহার করে থাকে যাতে করে ম্যালওয়্যার এনালাইসিস করা এবং ম্যালওয়্যার ডিটেক্ট করা আরো কষ্টসাধ্য হয়ে যায়। আর এই ধরনের Obfuscation কৌশলগুলোর মধ্যে একটি হলো **Virtualization-Based Obfuscation** । অর্থাৎ ম্যালওয়্যার এ এমন একটি টেকনলোজি থাকে যে যদি আপনি ম্যালওয়্যারটিকে আপনার ভার্চুয়াল মেশিন এ রান করেন তাহলে ম্যালওয়্যারটি তার কাজ বন্ধ রাখবে। এটিকে আবার **Anti-Vm** ও বলা হয়ে থাকে।

নিচে এ ধরনের টেকনলোজি ব্যবহার করা ম্যালওয়্যার এনালাইসিস করার রিসোর্স দেয়া হলোঃ

- Workshop: Analysis of Virtualization-based Obfuscation
  - writeup: https://synthesis.to/2021/10/21/vm_based_obfuscation.html
  - slides: https://synthesis.to/presentations/r2con2021-deobfuscation.pdf
  - video: https://www.youtube.com/watch?v=b6udPT79itk
  - code: https://github.com/mrphrazer/r2con2021_deobfuscation
- https://www.youtube.com/watch?v=PAG3M7mWT2c&t=13229s – a talk on reversing VMProtect

এবার শুধু রিসোর্স আর রিসোর্স এর পালাঃ

Courses

- https://www.begin.re/ – Reverse Engineering for beginners
- Reverse Engineering Malware 101 and 102 – by MalwareUnicorn
- https://github.com/mytechnotalent/Reverse-Engineering
- http://legend.octopuslabs.io/sample-page.html
- http://opensecuritytraining.info/Training.html
- https://samsclass.info/126/126_S17.shtml – Practical Malware Analysis
- Malware Analysis course (University of Cincinnati)
- Red/purple teaming: a malware development course by 0xPat
- My training: **malware_training_vol1** (work-in-progress)

YouTube channels

- Malware Analysis For Hedgehogs
- OALabs
- Colin's channel about malware
- DuMp-GuY TrIcKsTeR
- my channel

Books

- Practical Malware Analysis: A Hands-On Guide to Dissecting Malicious Software
- The Art of Computer Virus Research and Defense – Peter Szor
- "The "Ultimate"Anti-Debugging Reference" – by Peter Ferrie
- Malware Analyst's Cookbook and DVD: Tools and Techniques for Fighting Malicious Code
- Hacker Disassembling Uncovered – by Kris Kaspersky
- The Rootkit Arsenal: Escape and Evasion in the Dark Corners of the System
- Rootkits and Bootkits – by Alex Matrosov, Eugene Rodionov, and Sergey Bratus
- Windows System Programming (4th edition) – by Johnson M. Hart
- Gray Hat Python

আমি মনে করি যে আমার কাছে যা জ্ঞান আছে তা যথেষ্ট নয় কাউকে রোডম্যাপ দেয়ার জন্য। তবে নিজের কিছু এক্সপেরিয়েন্স এবং এই সাইবার সিকিউরিটি জগতে ক্যারিয়ার তৈরি করা কিছু ব্যক্তিদের থেকে তথ্য নিয়ে একটা সঠিক গাইডলাইন দেয়ার চেষ্টা করছি। ভুল ত্রুটি হলে ক্ষমাসুন্দর দৃষ্টিতে দেখবেন।

সমাপ্ত।