

Software Reverse Engineering

Software Reverse Engineering Quanto mai sarà difficile andare da Binario a Sorgente?

Gianfranco Gallizia

Linux Day 2022



Introduzione Disclaimer

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer

Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Strumenti Analisi statica Analisi dinam

Letture

Non sono un avvocato, né un giurista. Tutto quello che troverete in questa presentazione sono miei appunti sugli strumenti e sulle tecniche utilizzate per l'analisi di file binari che ho raccolto e sviluppato nel corso degli anni. Questa presentazione ha un puro scopo didattico ed illustrativo.

Non eseguite le tecniche descritte per fini illeciti ed in caso di dubbi consultate un professionista.



Introduzione

Che cos'è il Software Reverse Engineering (SRE)?

Software Reverse Engineering

Gianfranco Gallizia

Introduzione

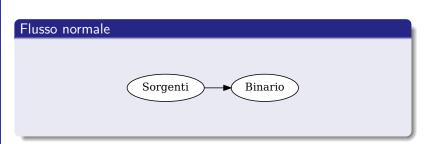
Disclaimer

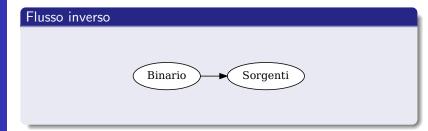
Analisi statica dinamica

Legalità del SI

Analisi statica

Letture







Introduzione Analisi statica e dinamica

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE?

Analisi statica e dinamica Legalità del SRE

Analisi statica

Letture

Analisi statica

Effettuata senza eseguire il software, guardando unicamente i bit su disco.

Analisi dinamica

Effettuata quando il software è in esecuzione guardando le attività ed il flusso di esecuzione.



Software

Reverse Engineering Gianfranco

Giantranco Gallizia

Introduzione
Disclaimer
Cos'è il SRE?

Analisi statica e dinamica Legalità del SRE

Analisi statio

Letture consigliate

Introduzione Analisi statica e dinamica

Analisi statica - pro e contro

- Pro:
 - Non si esegue il binario in esame. Bypassando le protezioni a runtime ed evitando di eseguire potenziali malware.
 - L'analisi può essere fatta su una macchina di architettura diversa dalla macchina bersaglio (ad esempio: analisi di un binario ARM su un pc amd64).
- Contro:
 - Non si esegue il binario in esame. Se il binario è compresso/offuscato/criptato/modificato a runtime non lo si vedrà.



Introduzione

Legalità del reverse engineering

Software Reverse Engineering

Gianfranco Gallizia

Disclaimer
Cos'è il SRE?
Analisi statica e
dinamica
Legalità del SRE

Analisi statica Analisi dinam

Letture consigliate

Quando è legale fare SRE nell'Unione Europea?

- Quando si ha l'esplicito consenso (meglio se scritto) del detentore dei diritti su software.
- Quando si effettua al fine di garantire l'interoperabilità del software con un altro software e/o con un altro sistema operativo/un'altra architettura. A patto che:
 - Si sia in possesso di regolare licenza d'uso per il software.
 - Non siano disponibili in altro modo le informazioni necessarie a garantire l'interoperabilità.
 - L'azione di SRE sia limitata alle sole parti necessarie a garantire l'interoperabilità.



Introduzione

Legalità del reverse engineering

Software Reverse Engineering

Gianfranco Gallizia

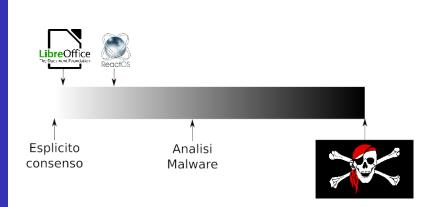
Introduzion

Disclaimer

dinamica

Strumen: Analisi stati

Letture





Strumenti del SRE

Di cosa ho bisogno?

Software Reverse Engineering

Gianfranco Gallizia

Introduzione
Disclaimer
Cos'è il SRE?
Analisi statica e
dinamica
Legalità del SRE

Strumenti

Analisi statica Analisi dinamica

Letture consigliate

Conoscenza

- Manuali, presentazioni, post su forum. Raccogliete quante più informazioni possibile.
- Pazienza
- Alcuni software
 - Editor esadecimale
 - Disassembler/Decompiler
 - System trace
 - Debugger
 - Virtual Machine/Emulatori
 - (Almeno) un linguaggio di programmazione/scripting



Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Analisi statica

Analisi dinam

Letture consigliate

Editor esadecimale

- Equivalente binario degli editor di testo.
- Usa la notazione esadecimale per rappresentare i byte del file (cifre da 0-9 e poi A=10, B=11, C=12, D=13, E=14, F=15). Due cifre esadecimali possono rappresentare i numeri da 00 a FF (255).
- Possono essere molto semplici (ad es. GHex) o con funzioni aggiuntive che semplificano l'analisi di strutture dati e la ricerca di pattern.
- Solitamente divisi in tre colonne: offset, dati in esadecimale, rappresentazione ASCII.



Software Reverse Engineering

Gianfranco Gallizia

Introduzion

Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Strument Analisi static

Letture consigliate

Editor esadecimale - Utilizzi

- Indagine preliminare:
 - È davvero un eseguibile?
 - Ci sono porzioni che contegono dati leggibili?
- Binary patching (sostituisco parte del binario per cambiarne il comportamento).



Software Reverse Engineering

Gianfranco Gallizia

Introduzione

Cos'è il SRE?

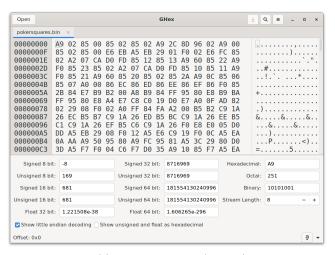
Analisi statica
dinamica

Legalità del SI

Strument

Analisi dinami

_etture consigliate



https://wiki.gnome.org/Apps/Ghex



Analisi statica - Editor esadecimale Hexplorer

Software Reverse Engineering

Gianfranco Gallizia

Introduzione

Disclaimer

Analisi statica

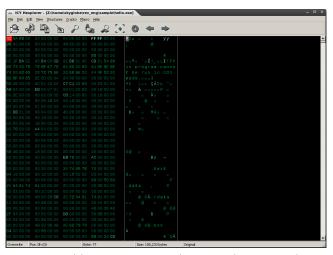
dinamica Legalità del S

Strument

Analisi statica

Analisi dinamica

Letture consigliate



https://sourceforge.net/projects/hexplorer/



Software Reverse Engineering

Gianfranco Gallizia

Introduzione

Disclaimer

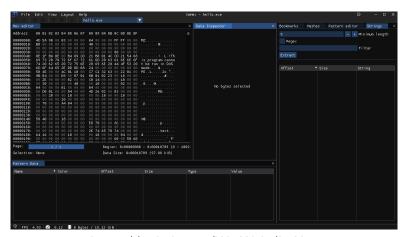
Analisi statica dinamica

Legalità del S

Strumer

Analisi statica Analisi dinamica

Letture



https://github.com/WerWolv/ImHex



Analisi statica - Disassembler

Software Reverse Engineering

Gianfranco Gallizia

Disclaimer
Cos'è il SRE?

Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Analisi statica

Letture consigliate

Disassembler

- Effettua l'operazione inversa di un assembler (passare da codice macchina a codice assembly).
- Strettamente dipendente dall'architettura bersaglio.
- In alcuni casi (CPU/microcontrollori legacy ad 8 bit) il disassemblato è di fatto il codice sorgente.
- Più frequentemente è il risultato finale di numerose trasformazioni effettuate dal compilatore per ottenere il codice macchina.



Analisi statica - Disassembler Esempi

Software Reverse Engineering

Analisi statica

Esempi

- objdump (Parte delle binutils)
- Radare2 (Suite di tools per SRE)
- Ghidra (Anche decompiler)



Analisi statica - Disassembler

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Strument Analisi statio

Letture consigliate

- Partito come tool per recuperare file da una partizione HFS+ si è evoluto fino a diventare un framework per il SRE.
- Ha un'interfaccia testuale scriptabile ed ha alcuni tool per compiti specifici: rasm2 assembler/disassembler standalone, rabin2 per l'analisi degli header binari, radiff2 per l'analisi delle differenze tra due binari.
- Supporta architetture multiple: amd64, ARM, x86, MIPS, PowerPC, Atmel AVR, Motorola 68k, MSIL, Java bytecode, Dalvik VM bytecode, WebAssembly (WASM) e altre ancora.
- È disponibile una GUI (laito).



Analisi statica - Disassembler

Radare2

Software Reverse Engineering

Gianfranco Gallizia

Introduzion

Disabilities

Analisi statica

Legalità del SI

Strumen

Analisi dinamica

Letture consigliate

```
skyglobe@gandalf: ~/rev eng/sample
                          print zoom view (see pz? for help)
                                        0x004014f9 mov dword [esp], str.libgcc_s
0x004014c0 mov dword [0x406070], 0
                                        0x00401500 call dword [sum.imp.KERNEL32.
0x004014ca imp sym. tmainCRTStartup
                                        0x00401506 sub esp. 4
                                        0x00401509 test eax, eax
            sym._atexit:
                                        0x0040150b je 0x401580
0x004014d0 sub esp. 0x1c
                                        0x0040150d mov ebx, eax
                                       0x0040150f mov dword [esp], str.libgcc_s
0x004014d3 mov eax, dword [esp + 0x20]
0x004014d7 mov dword [esp], eax
                                        0x00401516 call dword [sym.imp.KERNEL32.
0x004014da call sym.__onexit
                                        0x0040151c mov edi, dword [sym.imp.KERNE
                                        0x00401522 sub esp. 4
0x004014df test eax, eax
                                        0x00401525 mov dword [0x406028], eax
 x004014e1 sete al
                                        0x0040152a mov dword [esp + 4], str. re
 x004014e4 add esp. 0x1c
0x004014e7 movzx eax. al
                                        0x00401532 mov dword [esp], ebx
0x004014ea neg eax
0x004014ec
                                        0x00401537 sub esp. 8
                                        0x0040153a mov esi, eax
                                        0x0040153c mov dword [esp + 4], str.__de
                                        0x00401544 mov dword [esp], ebx
            sym.___gcc_register_frame: 0x00401547 call edi
0x004014f0 push ebp
                                        0x00401549 mov dword [0x403004], eax
0x004014f1 mov ebp. esp
                                        0x0040154e sub esp. 8
 0x004014c01>
```

https://www.radare.org



Analisi statica - Decompiler

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Analisi statio

Letture consigliate

Decompiler

- Effettua l'operazione inversa del compilatore.
- Molto più complesso di un disassembler (vi è più perdita di informazione da sorgente ad assembly che da assembly a codice macchina).
- Richiede molto più lavoro manuale da parte dell'operatore (leggasi: serve avere esperienza di programmazione).
- Solitamente il linguaggio di programmazione "ricostruito" dal binario è il C (con risultati "interessanti" nel caso in cui non sia così).



Analisi statica - Decompiler

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Analisi statica

Letture consigliate

Ghidra

- Progetto della NSA rilasciato come open source nel 2019.
- Scritto in Java e C++ (per le parti più computazionalmente intense).
- Può essere esteso con plug-in scritti in Java o in Python (grazie all'interprete Jython).
- Svolge sia funzioni di disassembler che di decompilatore (C/C++).
- Può interfacciarsi con un debugger (GDB in Linux, WinDbg in Windows).



Analisi statica - Decompiler

Software Reverse Engineering

Gianfranco Gallizia

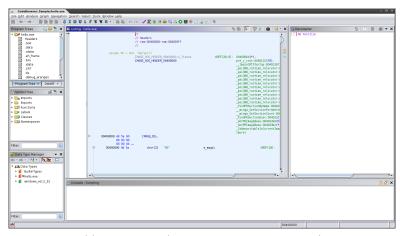
Introduzione Disclaimer

Analisi statica dinamica Legalità del SF

Strumenti Analisi statica

Analisi dinamid

_etture consigliate



https://github.com/NationalSecurityAgency/ghidra



Analisi dinamica - System trace

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Strumenti Analisi statica Analisi dinamica

Letture consigliate

System trace

- La system trace (o, più precisamente, system call trace) è una tecnica di analisi dinamica che intercetta le chiamate alle API del sistema operativo effettuate da un programma.
- Non c'è bisogno di modificare il programma analizzato, basta lanciarlo tramite un wrapper che istruirà il sistema operativo a fare rapporto al wrapper.
- Il programma tracciato può cercare di individuare il tracciante e modificare il suo comportamento di conseguenza.



Analisi dinamica - System trace Esempi

Software Reverse Engineering

Analisi dinamica

Esempi

- strace (Linux)
- ktrace (FreeBSD)
- Process Monitor (Windows Sysinternals)



Analisi dinamica - Debugger

Software Reverse Engineering

Gianfranco Gallizia

Introduzion Disclaimer Cos'è il SRE? Analisi statica e

Strumenti
Analisi statica
Analisi dinamica

Letture consigliate

Debugger

- Strumento usato dagli sviluppatori quando hanno bisogno di eseguire passo-passo il loro lavoro alla ricerca di problemi.
- Non è necessario avere i sorgenti di un programma per agganciare un debugger ad un processo: si riceverà un warning e l'analisi proseguirà a livello di disassemblato.



Analisi dinamica - Debugger Esempi

Software Reverse Engineering

Gianfranco Gallizia

Introduzion

.

Analisi statica

dinamica

Legalità del S

Analisi statica

Analisi dinamica

Letture

Esempi

- gdb
- x64dbg (Windows)



Analisi dinamica - Debugger x64dbg

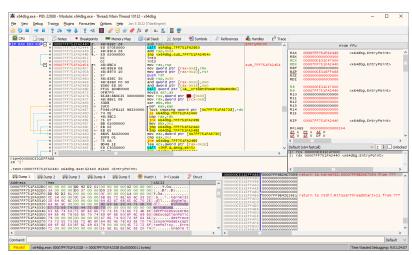
Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Analisi statica

Analisi dinamica



https://x64dbg.com



Analisi dinamica - Virtual Machine ed Emulatori

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Strumenti Analisi statica Analisi dinamica

Letture consigliate

Virtual Machine ed Emulatori

- Vi sono casi in cui non è possibile usare la stessa macchina per analizzare un programma ed eseguirlo:
 - Il programma è scritto per un'architettura diversa da quella della macchina utilizzata per l'analisi.
 - Il programma richiede hardware non presente sulla macchina utilizzata per l'analisi.
 - Non ci si fida del programma che si sta analizzando.
 - Si vuole analizzare il programma in un contesto il più pulito possibile.



Analisi dinamica - Virtual Machine ed Emulatori

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Strumenti Analisi statica Analisi dinamica

Letture consigliate

Virtual Machine ed Emulatori

- Le macchine virtuali (Virtual Machine, o VM) permettono di eseguire un programma in un contesto diverso da quello del sistema operativo della macchina ospitante:
 - Diverso kernel.
 - Diversa distribuzione.
 - Diverso set di pacchetti installati.
 - Diverso sistema operativo.



Analisi dinamica - Virtual Machine ed Emulatori Introduzione - cont.

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Strumenti Analisi statica Analisi dinamica

Letture consigliate

Virtual Machine ed Emulatori

- Quando i binari da analizzare non possono essere eseguiti sulla macchina corrente si può fare ricorso agli emulatori.
- Gli emulatori simulano hardware differente da quello presente sulla macchina che li sta eseguendo.



Analisi dinamica - Virtual Machine ed Emulatori

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Strumenti Analisi statica Analisi dinamica

Letture consigliate

QEMU

- QEMU è un emulatore e virtualizzatore open source.
- Può operare in autonomia oppure appoggiarsi ad un Hypervisor (solitamente KVM o Xen).
- Può operare in due modalità:
 - System mode emulation, in cui emula un'intera macchina.
 - User mode emulation, in cui emula solo la CPU e consente di eseguire un processo compilato per un'architettura su di un'altra.
- Supporta il protocollo di comunicazione remota di gdb e consente quindi di fare debugging di binari compilati per architetture differenti da quella della macchina su cui gira.



Analisi dinamica - Virtual Machine ed Emulatori

Software Reverse Engineering

Gianfranco Gallizia

Introduzione

Cos'è il SRE? Analisi statica e dinamica

Strumenti Analisi statica

Analisi dinamica





Software

Reverse Engineering

Giantranco Gallizia

Introduzion Disclaimer

Cos'è il SRE?

Analisi statica e
dinamica

Legalità del SRE

Analisi statica

Analisi dinam

Letture consigliate

Letture consigliate

Linguaggi di programmazione

Linguaggi di programmazione

- Brian Kernighan, Dennis M. Ritchie "The C Programming Language" ISBN 0-13-110362-8 272 pagine
- Steve Oualline "Practical C++ Programming" ISBN 978-0-596-00419-4 550 pagine
- Randall Hyde "The Art of Assembly Language" ISBN 978-1-59327-207-4 732 pagine



Letture consigliate

Sistemi operativi

Software Reverse Engineering

Gianfranco Gallizia

Introduzion

Disclaimer

Cos'è il SRE?

Analisi statica e
dinamica

Legalità del SRI

Analisi statica Analisi dinam

Letture consigliate

Sistemi operativi

- Brian Ward "How Linux Works" ISBN 978-1-71850-040-2
 464 pagine
- Michael Kerrisk "The Linux Programming Interface" ISBN 978-1-59327-220-3 1552 pagine
- Charles Petzold "Programming Windows" ISBN 978-1-57231-995-0 1479 pagine



Letture consigliate

Software Reverse Engineering

Gianfranco Gallizia

Introduzione Disclaimer Cos'è il SRE? Analisi statica e dinamica Legalità del SRE

Analisi statica Analisi dinam

Letture consigliate

Miscellanea

- Chris Eagle, Kara Nance "The Ghidra Book" ISBN 978-1-71850-102-7 608 pagine
- V. Anton Spraul "Think like a Programmer" ISBN 978-1-59327-424-5 256 pagine
- Brian Hook "Write Portable Code" ISBN 978-1-59327-056-8 272 pagine



Software Reverse Engineering

Gianfranco Gallizia

Introduzione

Cos'è il SRE? Analisi statica dinamica

dinamica Legalità del SR

Strument

Analisi statica Analisi dinamica

Letture consigliate

Grazie per l'attenzione

Live Demo con domande?