ANALYZING VOIP COMMUNICATION OVER H.323

Filip Karpíšek
Bachelor Degree Programme (3), FIT BUT
E-mail: xkarpi03@stud.fit.vutbr.cz

Supervised by: Petr Matoušek
E-mail: matousp@fit.vutbr.cz

Abstract: This paper describes background and design of a tool performing analysis and reconstruction of VoIP calls over H.323 standard. This tool has modular architecture and provides independent processing of input and export of metadata in XML format and audio data in both raw and converted (if possible) audiofiles.

Keywords: H.323, PER, ASN.1, VoIP analysis

1 ÚVOD

Tento příspěvek popisuje problematiku rekonstrukce komunikace postavené na signalizačních protokolech standardu H.323[1]. Tento standard je poměrně často používán a jeho analýza není triviální záležitostí. Rekonstrukce hovorů je součástí projektu SEC6NET zadaného Ministerstvem vnitra ČR.

V tomto článku představíme návrh nástroje, který by ze síťového provozu dokázal získat jednotlivé hovory a tyto exportovat v podobě audio a metadat. V současné době neexistuje volně dostupný nástroj, který by umožňoval zpracovat data VoIP zachycená ze síťového provozu – získat audio a metadata o hovorech, aby bylo možné jejich pozdější nezávislé zpracování (filtrování, řazení, vyhledávání podle určité položky, apod.).

Na závěr porovnáme výsledky analýzy dat VoIP s jinými nástroji, konkrétně s programy Wireshark a Cain & Abel.

2 NÁVRH A IMPLEMENTACE

Základní princip celého nástroje je postupné zpracovávání vstupního souboru paketů ve formátu pcap a podle informací získaných ze signalizačních zpráv dochází k detekci jednotlivých hovorů. Celý proces je jednoprůchodový, kdy používáme dynamický filtr. Pakety s hlasovými daty jsou přenášeny na předem neznámých portech, které se pro každý hovor liší. Proto je nezbytné analyzovat signalizační zprávy.

Pro implementaci bylo zvoleno prostředí jazyka Python, které umožňuje rychlé prototypování a obsahuje volně dostupné knihovny pro práci se síťovými daty.

2.1 ARCHITEKTURA APLIKACE

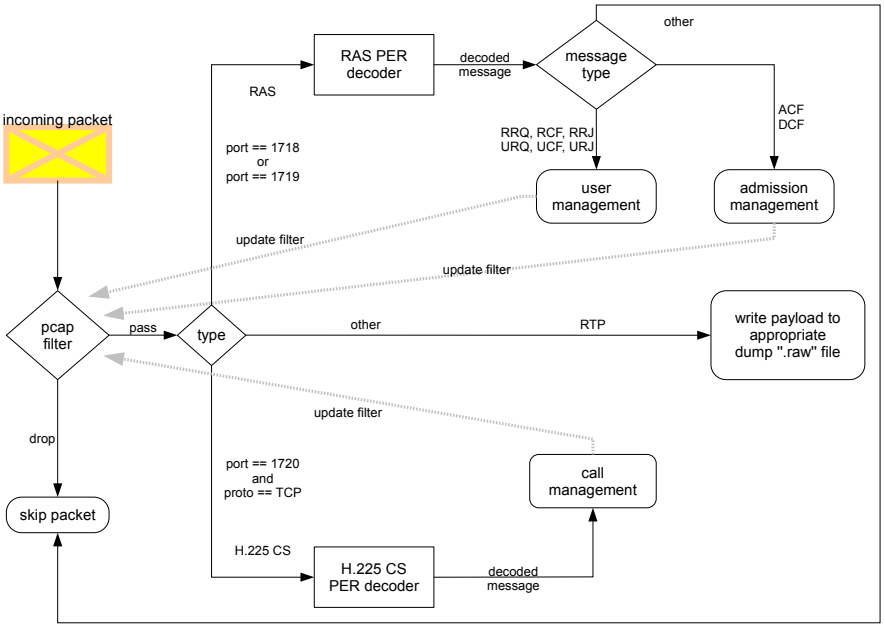
Obrázek 1 ukazuje celkovou architekturu nástroje. Příchozí paket prochází nejprve dynamickým filtrem, kde je na základě portu a případně adresy propuštěn nebo zahozen. Filtr propouští ve výchozím nastavení pakety s porty o hodnotách 1718 až 1720. Během zpracování pak dochází k dalším úpravám, viz kapitola 2.2.

Následuje klasifikace typu paketu. Paket je považován za zprávu H.225 CS, pokud je hodnota některého z portů transportní vrstvy (zdrojového či cílového) rovna 1720 a zároveň se jedná o TCP.

V případě hodnot portů 1718 nebo 1719 je paket klasifikován jako RAS. Všechny ostatní pakety (zbylé po filtraci) jsou považovány za RTP.

V případě RAS a CS zpráv následuje jejich dekódování a získání informací podstatných pro rekonstrukci. Tyto informace jsou předány do správce hovorů, který spravuje záznamy o hovorech. V některých případech dochází ke změně konfigurace dynamického filtru.

Zpracování RTP paketů je řízeno správcem hovorů, který vyhledá hovor, do kterého patří přijatý RTP paket, a zapíše obsah paketu do odpovídajícího výstupního souboru. Tímto způsobem jsou rekonstruována audiodata.



Obrázek 1: Architektura nástroje.

2.2 Komponenty systému

Načítání vstupních dat: Je využito knihovny pcapy, která umožňuje jak zachytávání ze síťového rozhraní, tak načítání souboru typu pcap.

Dynamický filtr: Jedná se o pcap-filter, kde jsou filtrační pravidla v textovém formátu. Ve výchozím nastavení propouští pakety na známých portech signalizačních protokolů (1718 – 1720). Při detekci nového hovoru postupně přibývají, při ukončení hovoru pak ubývají pravidla, díky kterým dochází k detekci paketů obsahujících hlasová data.

Dekodér PER: Zprávy standardu H.323 tvoří struktury v notaci ASN.1 kódované pomocí PER[3]. Toto kódování se vyznačuje velice efektivním využitím datového prostoru. Struktury ASN.1 navíc mohou obsahovat volitelné položky, které se nemusí vždy přenášet. To však přináší nevýhodu podstatně složitějšího dekódování. Pro zvolený implementační jazyk bohužel neexistuje volně dostupný dekodér, proto bylo nutné jej vytvořit.

Správce hovorů: Je jádrem celého nástroje. Zpracovává získané informace ze zpráv RAS a CS a udržuje aktuální informace v záznamech o hovorech. Pro každý hovor je uložen soubor informací sestávající se z informací nutných pro získání audiodat (adresy a porty hlasových toků) a

z dalších informací, které jsou podstatné pro praktické použití (začátek a konec hovoru, telefonní čísla nebo jiné identifikátory účastníků hovoru apod.).

Správce hovorů rovněž zpracovává pakety RTP, pro které je nutné vyhledat příslušnou signalizaci. Vyhledávání se děje na základě porovnávání adres a portů daného paketu s informacemi uloženými v záznamech o hovorech.

Export XML: Tato komponenta zajišťuje export získaných informací o hovorech do souboru XML, aby bylo možné pozdější nezávislé zpracování získaných dat. Exportér je přímo napojený na správce hovorů a transformuje získané informace do struktury XML.

3 SHRNUTÍ

V tabulce 1 jsou zaznamenány počty hovorů jednotlivých nástrojů jak pro data zachycená na páteřní síti VUT (CVIS), tak pro data vytvořená v laboratoři (LAB). Laboratorní data obsahují kompletní síťový provoz, data z reálného provozu pak jen část, většinou jednu stranu signalizačního toku, signalizace je tedy nekompletní. Při porovnání výstupů s programem Wireshark dochází k detekci stejného počtu hovorů, lze tedy hovořit o srovnatelném výsledku. Program Cain & Abel detekuje výrazně nižší počet hovorů (řádově jednotky procent) v porovnání jak s programem Wireshark, tak s vytvořeným nástrojem. K takovémuto rozdílu dochází hlavně u hovorů s nekompletní signalizací, program Cain & Abel tedy pravděpodobně využívá informace jen z některých zpráv, které v zachycených datech z reálné sítě chybí. Toto však není možné ověřit, neboť k programu Cain & Abel chybí dokumentace a program dál není vyvíjen.

data	CVIS 1	CVIS 2	LAB 1	LAB 2
Wireshark	300	201	4	2
Cain & Abel	20	5	4	2
vytvořený nástroj	300	201	4	2

Tabulka 1: Počet zachycených hovorů v různých souborech různými nástroji

Možné užití vyvinutého nástroje spočívá především v identifikaci jednotlivých hovorů, získávání statistik hovorů, účtování, apod. Nástroj může být rozšířen o další funkcionalitu, například export do HTML, kdy není nutný žádný speciální nástroj pro zobrazení, stačí běžný webový prohlížeč.

PODĚKOVÁNÍ

Tento příspěvek vznikl za podpory grantu MV a výzkumného záměru Moderní prostředky pro boj s kybernetickou kriminalitou na Internetu nové generace: VG20102015022.

REFERENCE

[1] ITU-T, Telecommunication standartization sector of ITU: H.323 v7 (12/2009)
Packet-based multimedia communications systems

[2] Vozňák, M.: Voice over IP. Skripta VŠB-TU, Ostrava, 2009

[3] Dubuisson, O.: ASN.1 Communication Between Heterogeneous Systems, Waltham, Massachusetts, Morgan Kaufmann 2000, ISBN 01-26333-61-0