SPA: A Smart Packet Analyzer for Network Traffic Analysis on Smartphones

Dilip Singh
Ajeenkya D Y Patil University, Pune

Ankit Kumar Singh
Ajeenkya D Y Patil University, Pune

Shabnam Sharma
iNurture Education Solutions, Bangalore

Chandan Prasad
Android Facilitator
Tata Strive, Pune

ABSTRACT
The intensifying growth of the internet and machinery whether its mobile or computer technology has brought countless good and proficient things for people such as E-commerce, E-mail, Cloud Computing, Data Sharing, Application and many more but there are also a dark and unseen sides of it such as Network Hacks, Computer hacks, Mobile Breach, Backdoors etc. In today's era, Cybercrime been one of the communal practices made by the computer specialists and is growing swiftly in numbers. Network monitoring performance of mobiles or computer are becoming increasingly significant for the security protection within today's organizational, ISP and carriers. There are the numerous tools which are available for network analysis, packet analysis such as Wireshark, Netresec, Termux. In this paper propose system is about an application idea through which user can analyze the network traffic, detect malicious packet, MAC spoofing, ARP spoofing, protocol filtering. Which will assist the user to identify the malicious packet that can cause damage to user's assets.

Keywords
Android, Network Monitoring, Packet Analysis, Termux, Wireshark.

1. INTRODUCTION
The usage of internet via mobile phones has already exceeded that of PC. In the meantime, the security problem of mobile phone and by mobile phone has become a general concern of the industry. Nowadays hackers are increasingly targeting mobile devices to perform various malicious through various mobile application, Phishing, Social Engineering.The pervasiveness of mobile devices mounts great pressure on to-day's network security infrastructures. Just like other desktop or web applications, mobile apps are supposed to be under the monitoring and safeguard of the security systems inside enterprise, ISP or carriers. Particularly, with the risk of potentially-harmful apps (PHAs) [24] [21] [22] on the growth, there is a robust demand for identifying them at the network level, using the anti-malware systems deployed by individual organizations or mobile carriers (through their Managed Security Services [23]). Smart analyzer will detect the malicious packet, applications in network by simply using packet filter technique. Smart analyzer provides user friendly environment to user, to revel the risk in an existing network.

1.1 Network Monitoring & Analysis
Network monitoring is a tough and challenging task that is a vital part of a Network Administrators job. Network Administrators are repetitively pushy to maintain smooth process of their networks. If a network were to be down even for a slight period of time, throughput within a company would decline, and in the case of public service divisions the ability to provide vital services would be compromised. In order to be upbeat rather than volatile, administrators need to monitor traffic movement and performance during the network and verify that security fissures do not occur within the network.

1.2 Packet Analyzer
A packet analyzer is also identified as a network analyzer, protocol analyzer or packet sniffer or for specific categories of networks, an Ethernet sniffer or wireless sniffer. It is a computer program that can capture and log traffic that passes over a digital network or portion of a network. As data streams flow transversely the network, the sniffer captures each packet and, if needed, decodes the packet's raw data, showing the values of various fields in the packet, and analyses its content.

1.3 Network Packet Analyzer
1.3.1 Wireshark
It is a free and open source packet analyzer. It is used for network troubleshooting, analysis, software and communications protocol development, and education. Formerly named Ethereal, the project was renamed Wireshark in May 2006 due to trademark issues.

1.3.1.1 The following are some of the many features Wireshark provides
i. Existing for UNIX and Windows.

ii. Capture live packet data from a network interface.

iii. Open files containing packet data captured with tcpdump/WinDump, Wireshark, and a number of other packet capture programs.

iv. Import packets from text files containing hex dumps of packet data.

v. Display packets with very detailed protocol information.

vi. Save packet data captured.

vii. Export some or all packets in a number of capture file formats.

viii. Filter packets on many criteria.

ix. Search for packets on many criteria.

1.3.2 Netresec Network Miner Packet Tracer
Netresec is an autonomous software vendor with emphasis on the network security field. This software is used for network forensics and analysis of network traffic. The most eminent product is NetworkMiner, which is accessible in a specialize0d as well as free open source version. They also develop and preserve other software tools, such as CapLoader (for big pcap files) and RawCap (a lightweight sniffer).

NetworkMiner is a Network Forensic Analysis Tool (NFAT) for Windows (but also works in Linux / Mac OS X /

FreeBSD). It makes it easy to perform advanced Network Traffic Analysis (NTA) by providing mined artifacts in an intuitive user interface. The way data is obtainable not only makes the analysis simpler, it also saves valuable time for the analyst or forensic investigator.

1.3.2.1 Features in Netresec Network Miner

i. Obtainable for Windows only.
ii. Capture live packet data from a network interface but not able to save the packets in any format for later use, it can analyses any packets file which is saved by any other software, then it can analyses the file only.
iii. Open files containing packet data captured. Wireshark, and a number of other packet capture programs.
iv. Export results to CSV / Excel / XML only
v. Display packets with protocol information.
vi. Can't filter packets on many criteria.
vii. No search for packets on many criteria.
viii. Colorize packet display based on filters but only on paid version.

1.4 Termux

Termux is an android terminal emulator and linux environment application that works directly without rooting android device. A nominal base system is installed inevitably, supplementary packages are offered to download using the Advance Package Tool (APT) package manager. It is a powerful terminal emulation with a wide-range of linux package collection. Pcap stands for "packet capture". A acquired file saved by Wireshark in the .pcap format. This file can be read by applications that recognize that format, such as tcpdump.

1.4.1 Features of Termux

i. User can edit the files using nano and vim.
ii. User can connect to server over SSH.
iii. It provides ability to compile code using gcc and clang.
iv. It also provides git cloning.
v. Works only on android with command interface.

1.5 Android Overview

Android is a mobile operating system developed by google, based on a modified version of the Linux kernel and other open source software[26]. The Linux kernel locates in the lowest layer of android system, while the Applications locate in the top layer. Android Applications can be written in Kotlin, java or C++ languages. The Android SDK tools compile code along with any data and resource files into an APK, an Android package, which is an archive file with an .apk suffix[25]. Every Android App runs in its own Linux process, and it secured by permissions and security sandbox. Each process has a unique user identifier (UID) and its own virtual machine (VM), so an App's code runs in isolation from other Apps. One android App comprises of four kinds of components: Activities, Services, Broadcast receivers and Content providers. Each kind has a different purpose and can by interconnect through Intent [27].

2. LITERATURE SURVEY

The goal of packet sniffing is to monitor network resources to detect anomalous behavior and misuse. This idea has been around for nearly 20 years but only recently has it seen a dramatic growth in popularity and incorporation into the overall information security infrastructure. Beginning in 1980, with James Anderson's inspiring paper [1], written for a government organization, familiarized the notion that audit trails contained vital information that could be treasured in tracking misuse and thoughtful user behavior. His work was the flinch of host-based intrusion detection and IDS in universal. In 1988, the Haystack project [2] released additional variety of intrusion detection for the US Air Force. This project shaped an IDS that analyzed audit data by equating it with defined patterns. In a conference, Crosby Marks, a prior Haystack Project team member and Haystack Labs employee said that," searching through this large amount of data for one specific misuse was equivalent to looking for a needle in a haystack." In 1990, Heberlein [3] introduced the awareness of network intrusion detection. Haystack Labs was the initial commercial vendor of IDS tools, with its Stalker line of host-based products. Nonetheless, commercial intrusion detection systems technologically advanced gradually during these years and only truly blossomed towards the latter half of the decade. In the last two decades, several network traffic classification techniques [4] [5] have been proposed to classify unknown classes. The primary one is Port Based Technique. It is a great technique for network traffic classification / identification. This practice includes a port, which is firstly registered in Internet Assign Number Authority (1ANA) [6]. Though, this system failed due to increase of Peer-to-Peer applications (P2P) in [7], which use dynamic port numbers. Dynamic port number means unregistered number with Internet Assign Number Authority (1ANA). Then second, one is Payload Based technique. This technique gives accurate results in network traffic classification. This practice is Deep Packet Inspection (DPI). Though, the problem is that it cannot be used for encrypted data network applications as several network applications use encrypted methods to protect data from detection. Therefore, this practice also failed due to use of encrypted flow of applications. Thereafter, the researchers proposed another method called Machine Learning Technique (ML) to categorize internet traffic as well as to know what type of applications flow in the network. Machine Learning Technique gives very capable accuracy results in network traffic classification. This practice is based on training and testing data sets to categorize unknown network classes. In paper [8] author defines the routing protocols using same Opponent software while they considered point to point throughput, querying delay and convergence time to compare the protocols. They both suggested EIGRP protocols for finest choice. In paper [9] they study performance of Virtual private Lan service network using Kerberos-enabled protocols (alternative authentication protocols) to degree the throughput value with respect to Normal VPLS network using Wireshark software IO graph. However, some other parameters like Delay, time factor, transmission efficiency is also important to precisely degree the concert of an authentication protocols in VPLS network. In paper [10] evaluated the routing protocols while they also measured combined routing protocols performance in IPv6 network using iperf software which measured the throughput, jitter and packet loss value in a same network's platform. In the paper [11], author scrutinized the performance of Ipv4 and Ipv6 when routing protocols have been utilized in both Ipv4 and Ipv6 virtual networks using GNS3.In this survey conclude that the Smart Analyzer application offers all those features which are not obtainable by existing applications as shown in fig.1 given below:

Tools / Features	WireShark	Network Miner	Netresec	Smart Analyzer
Packet Tracking	✓	✓	✓	✓
DDOS Detection	✓	✓	✓	✓
MAC Spoofing	✓	✓	✓	✓
ARP-Spoofing	✓	✓	✓	✓
Protocol Filter	✓	✓	✓	✓
SSH Tunneling	✓	✓	✓	✓
GUI Handeling	X	X	X	✓
Android	X	X	X	✓
Termux Support	X	X	X	✓
CLI	✓	✓	✓	✓
Result in PDF	X	X	X	✓
Result in Text	X	X	X	✓
Result in DOC	X	X	X	✓

Fig 1: Taxonomy Chart

3. PROPOSED SYSTEM

3.1 Phases

3.1.1 Input

In this Phase user interact with Smart Analyzer system over GUI and enters the input by selecting/choosing the various available scanning options such as TCP, UDP, MAC spoofing, ARP spoofing, Protocol filtering etc.

3.1.2 Smart Analysis

Smart Analyzer interprets the inputs provided by users into commands and then interacts with background running termux platform to execute these commands.

3.1.3 Execution

Commands provided by Smart Analyzer get manipulated by termux which has wireshark installed on it using Linux APT package manager. Wireshark manipulate these commands and returns the result to the termux and then termux returns to user over GUI interface provided by Smart Analyzer.

The steps of proposed methodology are depicted in figure 2, which consists of three major steps. These steps are: Smart Analyzer System User, Smart Analyzer, Termux and Wireshark

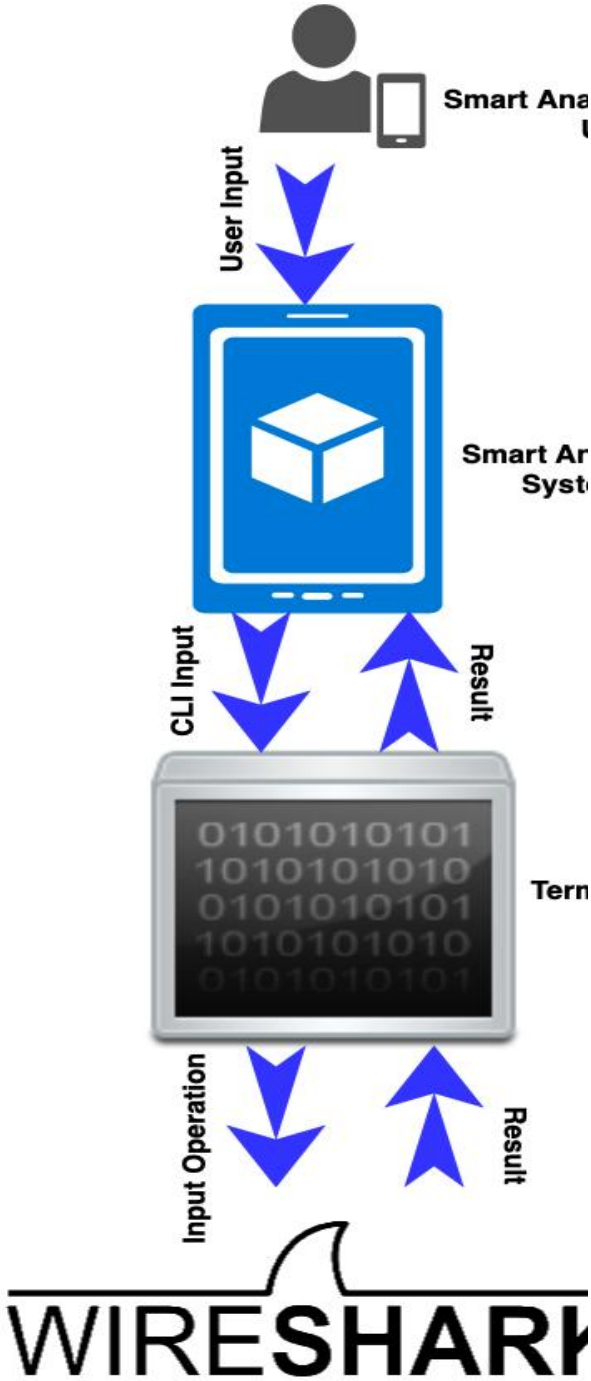

Fig 2: Proposed System Architecture

By implementing "Smart Analyzer" application user can able to improve the existing network security in a network. This application provides user friendly GUI to users which help them to analyze result in a better way. This application not only enhance network security but also provides additional features which are not obtainable by existing applications.

4. REFERENCES

[1] S. James P. Anderson, Computer security threat monitoring and surveillance", Technical report, Fort Washington, PA, April 1980.

[2] Stephen E. Smaha," Haystack: An intrusion detection system", In Proceedings of the Fourth Aerospace Computer Security Applications Conference, pages 37-44, December 1988.

[3] L. Todd Heberlein, Gihan V. Dias, Karl N. Levitt,Biswanath Mukherjee, Jeff Wood, and David Wolber, "A network security monitor", In Proceedings of the 1990 IEEE Symposium on Research in Security and Privacy, pages 296- 304, May 1990.

[4] Cao, Jie, et al. (2015): "Network Traffic Classification Using Feature Selection and Parameter Optimization.Journal of Communications 10.10.

[5] http://www.mathworks.com/help/stats/supervised-learningmachinelearning-workflow-and-algorithms.html.

[6] Arthur Callado, Carlos Kamienski, Geza Szabo, Balazs Peter GerYo, Judith Kelner,Stenio Fernandes, and Djamel Sadok.(2009): "A Survey on Internet Traffic Identification," IEEE Communications Survey tutorials, Vol. II, No. 3, pp. 37-52,Third Quarter 2009.

[7] Ian H. Witten and Eibe Frank (2005): Data Mining: Practical Machine Learning Tools and Techniques, 2th edition, Morgan Kaufmann Publishers, San Francisco, CA.

[8] S. Y. Jalali, S. Wani, M. Derwesh, "Qualitative Analysis and Performance Evaluation of RIP, IGRP, OSPF and EGRP Using OPNET" Research India Publications., Vol. 4, pp.389-396,2014. https://pdfs.semanticscholar.org/b616/f7b1a8e13f18b71998c557dc6f18d1fcba33.pdf.

[9] C. Fancy, L. M. M. Thanveer,"An evaluation of alternative protocols-based Virtual Private LAN Service (VPLS)" in IoT and Application (ICIOT), International Conference, Nagapattinam, India, May. 2017, pp. 1-6 (2017) https://ieeexplore.ieee.org/document/8073621/

[10] S.U. Masruroh, F. Robby, and N. Hakiem, "Performance Evaluation of Routing Protocols RIPng, OSPFv3, and EIGRP in an IPv6 Network" in International Conference on Informatics and Computing (ICIC), Mataram, Indonesia Oct. 2016, pp. 111-116 (2016). https://ieeexplore.ieee.org/document/7905699/

[11] D. R. Al-Ani, A. R. Al-Ani, "The performance of IPv4 and IPv6 in terms of Routing Protocols using GNS 3 Simulator" in 9th International Conference on Ambient Systems, Networks and Technologies, ANT-2018 and the 8th International Conference on Sustainable Energy Information Technology, SEIT 2018, May. 2018, pp. 1-6 (2018). https://dl.acm.org/citation.cfm?id=3223610

[12] Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems.

[13] Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

[14] Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems

[15] Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[16] Sannella, M. J. 1994 Constraint Satisfaction and Debugging for Interactive User Interfaces. Doctoral

Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.

[17] Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

[18] Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE.

[19] Y.T. Yu, M.F. Lau, "A comparison of MC/DC, MUMCUT and several other coverage criteria for logical decisions", Journal of Systems and Software, 2005, in press.

[20] Spector, A. Z. 1989. Achieving application requirements. In Distributed Systems, S. Mullender

[21] Kai Chen, Peng Wang, Yeonjoon Lee, XiaoFeng Wang, Nan Zhang, Heqing Huang, Wei Zou, and Peng Liu. 2015. Finding Unknown Malice in 10 Seconds: Mass Vetting for New Threats at the Google-Play Scale.. In USENIX Security Symposium. 659–674.

[22] Kai Chen, Xueqiang Wang, Yi Chen, Peng Wang, Yeonjoon Lee, XiaoFeng Wang, Bin Ma, Aohui Wang, Yingjun Zhang, and Wei Zou. 2016. Following devil's footprints: Cross-platform analysis of potentially harmful libraries on android and ios. In Security and Privacy (SP), 2016 IEEE Symposium on. IEEE, 357–376.

[23] Gartner. 2017. Managed Security Service Provider (MSSP).http://www.gartner.com/it-glossary/mssp-managed-security-service-provider/. (2017).

[24] Google. 2017. The Google Android Security Team's Classifications for Potentially Harmful Applications. https://static.googleusercontent.com/media/source.android.com/en//security/reports/Google Android Security PHA classifications.pdf. (2017).

[25] ApplicationFundamentals https://developer.android.com/guide/components/fundamentals.html

[26] Android(operation system)https://en.wikipedia.org/wiki/Android_(operating_system)

[27] Jice Wang and Hongqi Wu, "Android Inter-App Communication: Threats, Solutions, and Challenges" in arxiv.org on March 2018.https://arxiv.org/ftp/arxiv/papers/1803/1803.05039.pdf