International Journal of Security and Its Applications
Vol. 10, No. 9 (2016), pp.141-148
http://dx.doi.org/10.14257/ijsia.2016.10.9.15

Dynamic Security Policy Enforcement on Android

Matúš Vančo and Lukáš Aron

Brno University of Technology, Brno, Czech Republic
xvanco02@stud.fit.vutbr.cz, iaron@fit.vutbr.cz

Abstract

This work presentss the system for dynamic enforcement of access rights on Android. Each application will be repackaged by this system, so that the access to selected private data is restricted for the outer world. The system intercepts the system calls using Aurasium framework and adds an innovative approach of tracking the information flows from the privacy-sensitive sources using tainting mechanism without need of administrator rights. There has been designed file-level and data-level taint propagation and policy enforcement based on Android binder.

Keywords*: private data, Aurasium framework, operating system, system call, binder driver, Android security, policy enforcement, security policy*

1. Introduction

Android's fast growth of popularity has a lot of causes and consequences in terms of security. Growing number of applications, increasing a number of devices and growing level of integration have been interfered and influenced each other, implying increasing volume of the *private data*. People have put their trust in their devices and become more dependent on mobile technologies. This step is also necessary in order to using it for socialization, trading or entertainment. However, the private data is being used for the profit still more often during globalization, because it is the base for the knowledge-based business and targeted advertising. Even Google's free Android generates the significant part of its revenue just this way [1].

Since this asset is seized by many groups of people using illegal ways, Android has become the most assaulted mobile operating system facing the wave of malware infection, which is even more capable and stealthy, and can even establish a permanent presence on the device [2]. In order to address these challenges, Android includes permission model that protects access to sensitive resources. However, since permissions are overly broad and misunderstood, applications are provided with more access than they truly require. In particular, they are granted statically during install-time and so does not correspond to the actual use at the time. This implicates big vulnerability even if a certain application is not intended to misuse the private data because it can be exploited. There is also possibility to change the access rights after installation process manually, but who from the users do it?

Based on this insufficient built-in Android security and his later refinements, plenty of third-party frameworks seek to supplement overall security over the restrictions by the manufacturers. The current state of the art comprises several effective countermeasures to issues like coarse granularity of permissions, over-claim of permissions and permission escalation attack. However, most of these solutions are rather too complex and less straightforward. They replace the whole Android permission model, or tries to track the information flow on the level of operating system which requires the rooted device. In contrast, there is Aurasium framework, which automatically repackage and harden chosen applications, interposing the sandboxing code in the applications themselves. Nevertheless, it is robust enough to interpose almost all types of interactions between the application and the operating system.

ISSN: 1738-9976 IJSIA

The aim of this work is to present the system, which utilizes the Aurasium framework and restrict the access of private data outside the device through the selected applications. In contrast to original Aurasium framework, this work focuses more on the real asset, the private data, and especially on the high usability and easy deployment. The solution is focused on tracking the information flows from the privacy-sensitive sources to the system sink where they aim to leave the system.

2. Related Works

Implementations of countermeasures to problems have been divided into several categories according to respective addressed issues [5]. Compared implementations address the three of six mentioned issues – *Coarse Granularity of Permissions* (CGP), *Over-claim of Permissions* (OCP) and *Permission Escalation Attack* (PEA). These categories have been further divided to provide more detailed overview. Resulting categories and their overlapping have been illustrated in Figure 1.

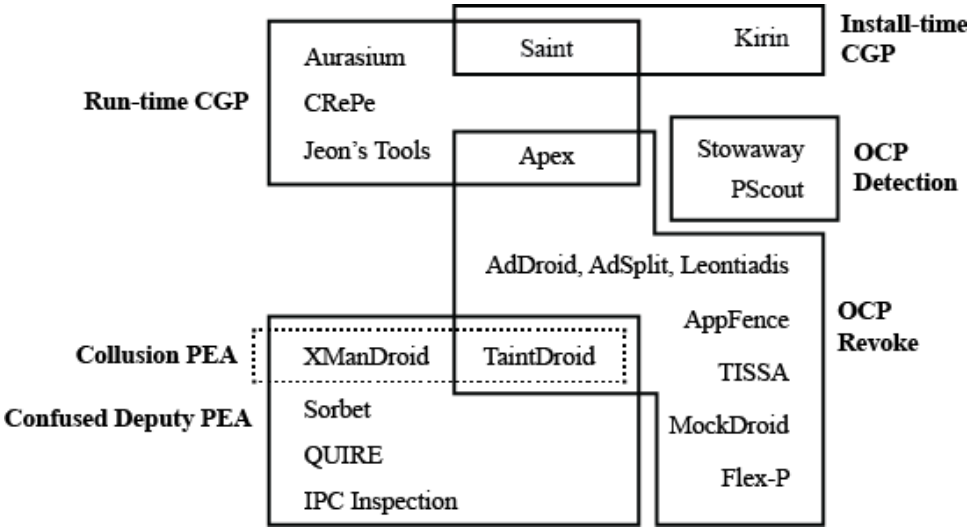

Figure 1. Countermeasures According to Addressed Issue

Enforcement can be performed during install-time or during run-time. Install-time enforcement means that applications that do not pass the established policy are prevented from being installed. On the other hand, run-time enforcement enables installation itself, but dynamically prohibits access to protected system resources or other application components. In that case, some applications exit unexpectedly with an exception, some can continue their activity with false data provided, some can prohibit starting an application in certain context. Also information sources for establishing the policy rules are different. Some applications use policies according to user wishes, some analyze all existing permissions in the system or content of intents and creates rules accordingly to prevent various attack, some are defined statically. In addition, each solution has been working on the different layer of the Android OS and only minority of them does not require administrator privileges to the system. Also the information source for the decision making about policy enforcement differs. It can be an author of application, a user of application or there are used existing security policies to decide whether allow or block certain action.

2.1. Aurasium Framework

Aurasium project has been developed in 2012 as a project at the University of Cambridge, UK. Aurasium use repackaging mechanism, wrapping around the DVM under which the Android applications run, with monitoring code. It does not require rooted Android device. To attach sandboxing code, Aurarium exploits Android's unique application architecture of mixed Java and native code execution and introduces *libc* interposition code. Because of this, Aurasium is capable to mediate almost all types of interactions between the application and the Android OS. This project consists of three parts – automated repackaging system written in Python programming language named "pyAPKRewriter", sandboxing code included in "ApkMonitor" application and "Aurasium's Security Manager" (ASM) application enabling central handling of policy decision of all repackaged application on the device. [2]

Starting with sandboxing code, the top layer of the framework is written in Java. The aim is to create a well-documented easy-to-use abstraction layer upon cumbersome native layer of the framework. The upper layer creates interface for other possible programs and delegates all requests to the low-level part of the framework implemented in native C++ code. This layer consists of few shared objects that do all the real work, such as communication with the Dalvik VM or establishing the mechanism for IPC communication.

The second part of Aurasium, the repackaging Python script utilizes the previously mentioned sandboxing code and deploy it to Android's installation package – APK file. APK file is similar to Java JAR archive and contains Android Manifest file with declared permissions, application logic in the form of dex bytecode, compiled XML resources and native libraries. Each application package is also signed with authorship information. Besides the sandboxing code, Aurasium has to include also several additional parts to APK in order to ensure the functionality.The last part of the Aurasium is called *Aurasium Security Manager* (ASM). ASM handles the policy decisions centrally, so that all repackaged applications can be maintained at one place. Security policy is based on decision of application or user. Application decision works transparently without user interaction, while the user decision is consented by dialog window and can be remembered and used by default during next occurrence.

3. Problem Analysis

Android security has been built upon fundamental security concepts of the operating systems themselves. Android's Permission Label Model (PAM) has been built upon Linux security mechanisms including Mandatory Access Control (MAC) mechanism and Principle of Least Privilege (PLP) [3]. Security enforcement using MAC uses two types of permissions – granted permissions (used or requested permissions) and required permissions (access permissions). Granted permissions are manifested during installation and are inherited by all of the application's components. On the other hand, required permissions are usually created by the developer to protect their important components. Required permissions are always assigned to application components separately. When the application is started, the launcher component is invoked and the other components are called subsequently from the same or the other application or system. The mechanism of the passing to another component and access control is transparent in both situations. Communication between components is based on Linux Inter-component Communication (ICC), because each application runs in the separate process. It uses message passing, where messages contain data with the required action and are called intents ("ACTION_SEND", "ACTION_VIEW", *etc*.) [4]. Android maintains this communication using Reference Monitor (RM), which is part of the Android OS Middleware.

3.1. Android Binder

In order to perform the required mediation, the part of Android middleware called the Binder needs to be rewritten. The Binder was originally developed under the name *OpenBinder* by *Be Inc.* and later under *Palm Inc.* and provides high-level abstraction on top of traditional modern operating system services including the facility to provide bindings to functions and data from one execution environment to another [6]. In Android, OpenBinder is customized to provide Inter-component Communication as described before. All interposition code needs to be placed in the suitable position in the original Binder implementation. Therefore, there is important to understand the concepts and to analyze the architecture of this part of system.

The communication between two processes is ensured using Binder Objects (BO), which are instances of classes that implement ioctl-based Binder interface. The most important operation which is declared in this interface is "transact (int code, Parcel data, Parcel reply, int flags)". The corresponding callback method in the Binder object is called "onTransact". The interface can be further extended by additional business operations using Android Interface Definition Language (AIDL). Each BO uses local and global identifier. The local ID is unique in the process and the global ID is created when the BO is passed to another process using Binder Driver (BD). The BD then works like network switch and persists the mapping from local ID to global ID in the table structure and translate it transparently, similarly than the mapping using ARP protocol. The Binder framework communication uses the client-server model. However, the process can implement the server as well as client, so the communication can be still bidirectional. The Binder Client (BC) invokes an operation on a remote Binder object called Binder Transaction (BT), which may involve sending or receiving data over the Binder Protocol. The communication is performed indirectly using Binder Driver. In the Android, the Binder Driver is exposed via "/dev/binder" file and simple API based on "open", "release", "poll", "mmap", "flush" and "ioctl" operations. Most communication happens via "ioctl(int fd, unsigned long request, ...)" method. The first parameter is the file descriptor number which identifies currently opened file and is used in "/proc/<pid>/fd/<fd>" file. The second parameter specifies the request. In fact, most communication happens via "ioctl(binderFD, BINDER_WRITE_READ, &bwd)" operation, where the "binderFD" is used to access the "/dev/binder" file and the "bwd" structure ensures communication using read and write buffer. Illustration of this structure is shown in Figure 2.

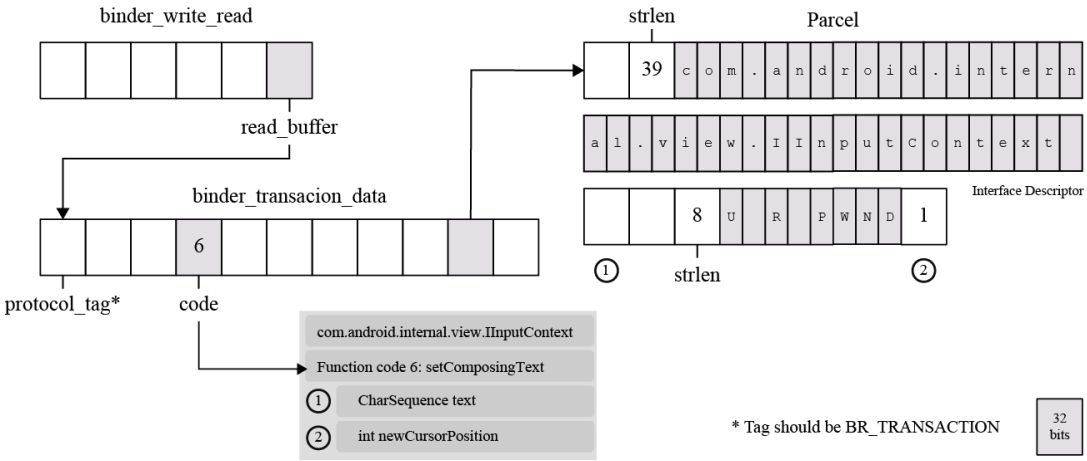

Figure 2. Binder Write-read Structure[1]

[1] Inspired by [8]

The "write_buffer" contains a series of commands for the driver to perform, while the "read_buffer" contains commands for the BO in user-space. The commands for driver are called Binder Call (BC) commands and the commands for the BO are called Binder Return (BR) commands. Each command is couple (operation code, data). The Binder Transaction is a passing data from the client to the service, while the Binder Reply is a passing data from the service back to the client. This is shown in Figure 3.

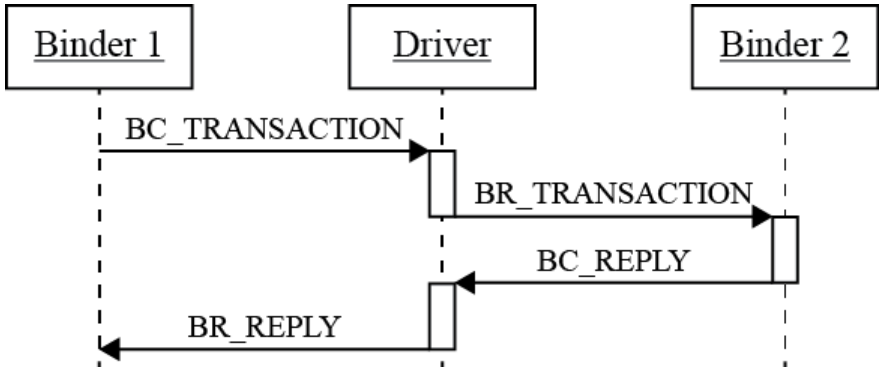

Figure 3. Binder Driven Interaction [6]

The whole Binder framework mechanism is transparent for the Android developer, since the Binder Transaction is performed as a local function call using so-called thread migration. This is ensured by the proxies and stubs, which are auto-generated helper classes from the AIDL files [7]. The proxy is the helper class which transforms Java code to low-level commands for the Binder Driver. The stub works in reverse to proxy and automatically parses and performs read commands on the service side. Since the Binder Driver is implemented on the low layer using C language, there has to be mechanism for encapsulation of high-level Java objects. This is ensured by "Parcel" container and corresponding "Parcelable" interface. A procedure for converting this higher-lever applications data structures into parcels is called marshalling. The marshalling as well as unmarshalling are also in the responsibility of the proxies and stubs.

4. Data Tainting and Restriction

Design of the system is based on previous analysis and various experiments, which were focused on the Android system behavior. Design consists of three parts – design of architecture and principle of application, design of data structures and design of configuration. Design of architecture can be further devided into two parts – desing of tainting mechanism and design of restriction.

Starting with the overall architecture and tainting principle, the tainting is based on the principle used in TaintDroid architecture. In order to perform real memory-level tainting, there has to be tracked each atomic memory transfer, which from a developer's point of view means each assignment to a variable. This can be done only through monitoring of instructions on the level of machine. In TaintDroid, there are monitored instructions on the level of virtual machines, because all possibly harmful applications are run under Dalvik virtual machine. TaintDroid uses *Virtual Taint Map* (VTM), which mirrors the address space, but does not contain the content of memory. It represents the division of memory into protected and public part. Before tainting process, the tainted files are marked in VTM. Then, every copying of memory invokes copying of blocks in VTM. Since the applications, which run on separate DVM can also exchange data, TaintDroid introduces message-level tainting as well.

In this project, there has been designed and integrated two granularities of taint propagation – file-level tainting and data-level tainting. The message-level tainting (between components) principle from TaintDroid is used only for final policy

enforcement (restriction), because Aurasium intercepts only single applications and does not have possibility to monitor the unhardened ones. File-level tainting and data-level tainting uses the previously mentioned concept of VTM, but it is stored in higer-level abstract data structure and file instead of VTM.

File-level tainting between memory and the OS's file system can be performed in full scope, because Aurasium can fully intercept this communication using system calls "fopen", "open", "write" and "read". Fuction "fopen" is used for obtaining the opening mode. This is used for *tainting customization* (TA). If the untainted memory is written to tainted file in append mode, the files remains tainted, but if it is written in read mode, the file becomes untainted. The "open" and "read" calls are used for tainting the memory blocks as well as new files. The data in memory read from tainted file are marked similarily and the files, which are read from tainted memory blocks becomes tainted too. However, data in memory are also directly propagated.

Since the Aurasium can intercept only specific places (system calls) and not instruction itself, it is impossible to implement full-scope memory-level tainting as is introduced by TaintDroid. This is replaced by the newly designed data-level tainting concept. This concept together with the file-level tainting is depicted in Figure 4. When the data are read from the file, content of data is read and tagged using hash function which assigns a unique number. This tag, together with the size of block is used during the writing unknown memory block into file. Each unknown memory block is tested with respect to any existing hash and marked as tainted if the hash matches. Subsequently, the file is marked as tainted as well.

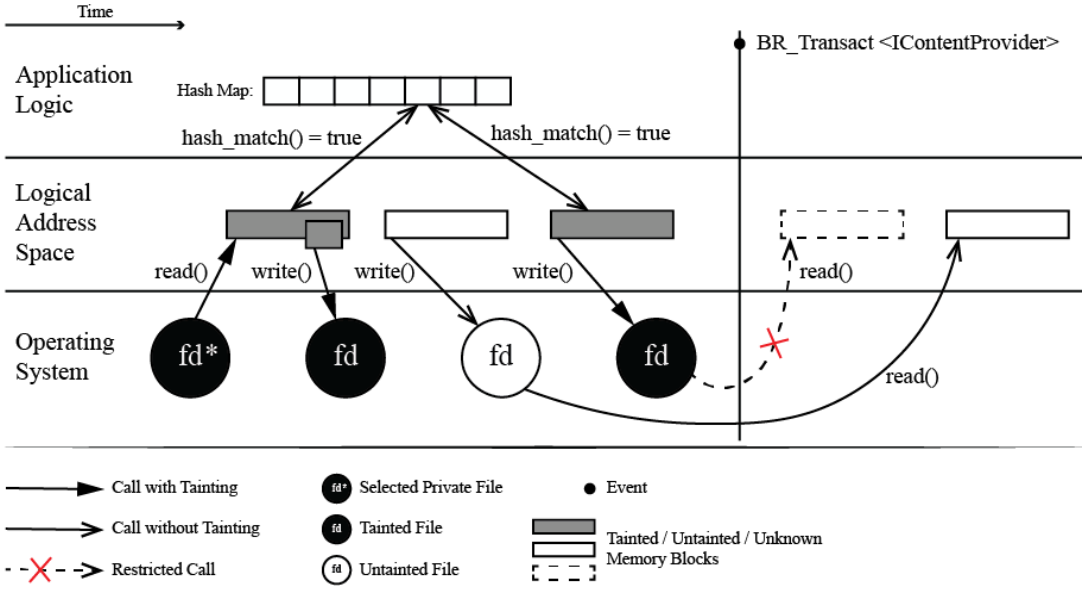

Figure 4. Design of Architecture

The final policy enforcement is performed using the interception of "ioctl" call. Specifically, when the "BR_TRANSACTION" command which contains destination component ContentProvider is read, all the "read" calls for the tainted files are in the mode of restriction.

The project is designed to secure the user-selected files or folders as a entity, which are intended to be invariable like images, pictures or videos. Documents that are often changed can be restricted for opening, or there can be assigned unique rights for opening to hardened application and the files are encrypted for other applications (reverse mode). In this reverse mode, data are protected with unhardened applications and uncovered and possibly exploited by the hardened application. The reverse mode is designed as optional and may not be implemented. Further extension is finer granularity of data-tainting.

Unknown memory block which is being written to file is compared against the tainted memory blocks which are smaller than the unknown memory block, and the memory block which is being read from file has been divided into smaller units with separate hash.

Data structure which works as TaintDroid's VTM is designed as simple array of memory blocks, which are the interconnection part between the file system level and application logic performing described data-tainting. From designed perspective, each memory block is considered as tainted or untainted. Application can store only tainted data and other will be implicit. Initially, only the files are marked as tainted and during tainting process, the other files and new memory blocks are added. Each memory block can have only one source file, one hash tag, but a lot of destination files to which this data is written. Due to Tainting Customization (TA), also the file modes need to be stored, because "read" and "write" functions does not dispose with this information inside passed arguments.

Regarding the configuration possibilities, the main purpose is to allow the selection of private files and folders. Since the two previously-mentioned tainting mechanisms are preferred in different situations, the setting of selected mechanism is also appropriate. Hash-only tainting can be used in situations where the low memory consumption is important, while the content-based scanning in the case of files which need better protection. In some cases, there is also a need to disable the tainting completely due to performance slowdown. The restriction can be performed explicitly as well as dynamically. Explicit (static) restriction can be used in situations which require protection from theft or unauthorized users. One exemplary utilization includes parental control. Dynamic enforcement can be realized using confirmation screen and is useful for its flexibility. It is also appropriate to consider the configuration of permanent restriction where the selected protected files can not be even read by hardened application. The configuration settings can be assigned to each application separately or centrally. GUI can be also resolved as a single central application which communicates which hardened applications or as injected screen in each application. For the purposes of this project, the first variant is chosen. The advantage is an easier configuration for multiple applications at once, usability and better overview of the whole configuration in one place. However, there is security issue related to the communication between configuration application and the hardened ones. The communication is designed via configuration file for its simplicity. This file should be secured to ensure mainly the data integrity. Returning back, the usage with the permanent restriction, especially as a parental control, requires the password protection of configuration application itself.

5. Conclusion

The aim of this work was to provide the solution for securing the user-selected private data of chosen applications with the sandboxing mechanism. In the first phase, inter-process communication and existing frameworks, which are capable of intercepting communication between the application and the operating system on the level of system calls, were explored. Subsequently, the possibilities and the code of the one of the compared frameworks – Aurasium framework – are investigated. From this analysis, the framework has been modified to be able to build and proper run on the latest required Android version, and there were conducted experiments in order to propose and realize the tainting mechanism as well as the security policy enforcement.

The system performs the file-based tainting of selected files using hash calculation and content-based tainting using division of file content into "small blocks". The restriction is implemented intercepting the communication with an external application using "ContentProvider" interface and via falsifying the file content. There is also developed restriction based on falsifying the whole files on the file system. Application introduces communication with the configuration application utilizing the configuration file, which is

similar than low-level Android binder communication. Configuration options include the type of scanning, type of restriction, method of intercepting or static restriction of file opening. The system has been developed and tested in three different types of environment – in mocking C++ environment, in selected hardened test applications and in Aurasium repackaging system. The C++ environment is prepared on Linux machine with several test scenarios, the hardened applications are mostly tested on the real Android device using the Android IDE and the Aurasium repackaging system is tested by replacing the monitoring code in the original *ApkMonitor* package.

The next steps are the code optimization, design, and implementation of the secure communication between configuration activity and the hardened applications, addressing and handling the file permissions or the implementation of several proposed extensions. The next work can be also aimed at developing the automated configuration and risk analysis, unique handling the files according to the type of media (image, video, text files), expanding the configuration options, focusing on the possibilities of the Linux core or further testing, experimenting and research.

Acknowledgments

The work was supported by the IT4Innovations Excellence in Science project LQ1602 and the internal BUT project FIT-S-14-2486.

References

[1] J. RosenBlatt, "Google's Android Generates 31 Billion Dollars Revenue, Oracle Says", **(2016)**.
[2] R. Xu, H. Saïdi and R. Anderson, "Aurasium: Practical Policy Enforcement for Android Applications", Proceedings of the 21st USENIX conference …, **(2012)**, p. 27.
[3] A. Silberschatz, P. B. Galvin, G. Gagne and A. Silberschatz, "Operating system concepts", 9th ed., Addison-Wesley Reading, vol. 4, **(1998)**.
[4] Android, "Android Developers", **(2016)**.
[5] Z. Fang, W. Han and Y. Li, "Permission based Android security: Issues and countermeasures", Computers & Security, vol. 43, **(2014)**, pp. 205-218.
[6] T. Schreiber, "Android binder", Ruhr-Universität Bochum, **(2011)**.
[7] A. Gargenta, "Deep dive into Android IPC/Binder framework", in AnDevCon: The Android Developer Conference, **(2012)**.
[8] N. Aetenstein and I. Revivo, "Man in the Binder: He Who Controls IPC, Controls the Droid", in Europe BlackHat Conf, **(2014)**.

Authors

Matúš Vančo, was born in 1991, Trnava, Slovak Republic. He received the B.S degree in information technology from Brno University of Technology (BUT) in Czech Republic, 2012. In 2013, he received M.S. degree in computer science at University of Glamorgan in UK. Now, he is a postgraduate student at BUT, dealing with Android development and IT security.

Lukas Aron, was born in 1987, Prague, Czech Republic. He received the B.S and also M.S. degree in information technology from Brno University of Technology (BUT) in Czech Republic. Now he is student of Ph. D at BUT, dealing with IT security. Besides the university he is senior software developer at AVG Technologies.

