PINPOINT: Efficient and Effective Resource Isolation for Mobile Security and Privacy

Paul Ratazzi
Ashok Bommisetti
Nian Ji
Wenliang Du

Available at: https://works.bepress.com/ratazzi/7/
PINPOINT: Efficient and Effective Resource Isolation for Mobile Security and Privacy

Paul Ratazzi∗†, Ashok Bommisetti†, Nian Ji† and Wenliang Du†
∗Information Directorate, Air Force Research Laboratory, Rome, NY
†Dept. of Electrical Engineering & Computer Science, Syracuse University, Syracuse, NY

Abstract—Virtualization is frequently used to isolate untrusted processes and control their access to sensitive resources. However, isolation usually carries a price in terms of less resource sharing and reduced inter-process communication. In an open architecture such as Android, this price and its impact on performance, usability, and transparency must be carefully considered. Although previous efforts in developing general-purpose isolation solutions have shown that some of these negative side-effects can be mitigated, doing so involves overcoming significant design challenges by incorporating numerous additional platform complexities not directly related to improved security. Thus, the general purpose solutions become inefficient and burdensome if the end-user has only specific security goals.

In this paper, we present PINPOINT, a resource isolation strategy that forgoes general-purpose solutions in favor of a “building block” approach that addresses specific end-user security goals. PINPOINT embodies the concept of Linux Namespace lightweight isolation, but does so in the Android Framework by guiding the security designer towards isolation points that are contextually close to the resource(s) that need to be isolated. This strategy allows the rest of the Framework to function fully as intended, transparently. We demonstrate our strategy with a case study on Android System Services, and show four applications of PINPOINTed system services functioning with unmodified market apps. Our evaluation results show that practical security and privacy advantages can be gained using our approach, without inducing the problematic side-effects that other general-purpose designs must address.

I. INTRODUCTION

Over the decade since its introduction, Android has been a stunning success, eclipsing the market share of every other mobile operating system by a huge margin, and now shipping on well over 1 billion new devices annually [16]. This growth, however, has not been without its pains. By one recent measure, 97% of today’s mobile malware targets the Android operating system [13]. Commensurate with this trend, end-users are increasingly concerned about privacy and protecting personal information. Unsurprisingly though, the typical end-user possesses little or none of the specialized technical expertise necessary to fully understand the security implications of installing apps, granting permissions, entering sensitive data, etc. As a result, most users have a hard time using currently-available security indicators to identify which apps they should trust, and which they should not [6]. Faced with this lack of knowledge and confusing choices, many become complacent or careless in performing what amounts to critical system administration tasks.

Even though successive releases of Android continue to enhance and improve security [2], balancing security with usability has proven difficult. For example, Android 4.3’s App Ops feature for selective permission granting was hidden from end-users in version 4.4.2, apparently due to usability concerns. Because developers cannot anticipate the endless security configurations App Ops makes possible, many apps failed to function or simply crashed when their permissions were selectively revoked by the user [10]. Although this functionality is still obtainable using third-party apps, most uninformed users are unlikely to make the additional effort of activating and understanding these features at the level of technical detail necessary. As a result, even the latest releases of Android 5.0 do little to help end-users protect themselves from software they choose to install.

While many users may not fully understand the technical aspects of security architectures, permissions, access control mechanisms, or measuring trust, most have no trouble articulating which high level objects, resources or capabilities they are most concerned with. For example, it’s common to find users worried about how some apps might misuse location, sensitive data such as personal contacts, or personally-identifiable information (PII) like phone number or International Mobile Station Equipment Identity (IMEI). In response to this, numerous solutions to address these concerns have been proposed, and many of these involve some form of virtualization combined with access control to isolate untrusted applications.

Although every approach to isolation has its own unique strengths and weaknesses, all include trade-offs in terms of sharing and communication. In Android’s open architecture, where resource sharing and inter-process communication (IPC) are fundamental to the platform’s basic operation and usability, careful attention must be paid to fully understanding how a particular isolation boundary impacts the system’s functionality and performance. If this trade-off is not considered at the outset of a design, significant performance, usability, and functionality issues can arise. Countering these negative side-effects requires designers to overcome challenging system problems, typically resulting in substantial modifications to the operating system, and significant second-order complexities not directly related to the initial security goals. These problems are especially prevalent in general-purpose designs that attempt to provide isolation containers for entire apps or virtual phones, without the benefit of a priori knowledge of specific threat(s) or end-user security goals.

In this paper, we present PINPOINT, a resource isolation strategy that forgoes general-purpose solutions in favor of a “building block” approach that addresses specific end-user security goals. By addressing stated security goals and no more,
PINPOINT yields an effective result using only the minimum amount of isolation. This helps minimize or eliminate the negative side-effects that are sure to emerge when large parts of Android’s open architecture are subject to isolation. The PINPOINT concept and its scope of applicability in Android are introduced in Section II. Section III describes a case study whereby we implemented the PINPOINT concept as a lightweight hypervisor\(^1\) within Android’s Context Manager in order to make possible isolation of any system service. Sections IV and V contain the implementation details and evaluation results of isolating four Android system services using this hypervisor. We summarize related work in Section VI, future directions in Section VII, and conclude the paper in Section VIII.

II. Motivations & Concept

Our idea for PINPOINT originated from the realization that much of the time and effort devoted to implementing current isolation architectures (summarized in Section VI) is spent on overcoming the negative side-effects on the Android system introduced by the chosen point of isolation. Most times, eliminating or mitigating these side effects requires challenging and far-reaching operating system modifications. Even when these challenges are met, we found ourselves unsure whether end-users would tolerate the remaining reductions in functionality, convenience, and performance. Furthermore, we thought it to be unlikely that these complex and less-usable designs would be adopted by Google or original equipment manufacturers (OEM), since Android user experience and its open architecture are paramount.

Nonetheless, we were intrigued by these designs’ use of Linux Namespace lightweight isolation and were inspired to further explore the possibilities. Our conclusion is that for certain security scenarios, lightweight isolation using Linux Namespaces has many interesting advantages over heavy-weight forms of isolation, such as virtual machines. Through a systematic analysis of Linux Namespaces, we identified six key traits that have value to our goals of providing effective yet efficient security. These traits are summarized in Table I.

On the other hand, we believe that the authors’ direct use of Linux Namespaces as the point of isolation underlying the Android Framework breaks several basic assumptions of Android’s open design. Hence, this choice of kernel-level isolation is the root cause of many of the complex system problems and negative side-effects they encountered. Although choosing an existing low-level mechanism as the foundation for isolation enabled their general-purpose isolation containers, there are many times when the cost of such comprehensive isolation is not worth it, especially when the end-user’s security goals are relatively simple. For example, preventing a single untrusted app from accessing the device’s IMEI. Our work fills this gap by providing an approach that moves the point of isolation as close as possible to the object(s) needing isolation, based on the stated security goals. In essence, we strive to realize the benefits of lightweight isolation in Android by identifying strategic locations where Linux Namespace concepts can be implemented. As we will show, the result is a simpler implementation free of problems and far-flung platform side-effects. When specific security goals are taken into account, the result can be just as effective as general-purpose solutions.

Figure 1 depicts our general, high-level concept. Given a specific security goal that relates to App 2’s interaction with objects A and B, our goal is to place the point of isolation (i.e., the hypervisor) at a strategic location that enables virtualization of all these objects such that App 1 and App 2 see different instances of each, while everything else about the system is common and unmodified. When trusted App 1 requests A and C (1), the hypervisor returns instances of A and C (2). On the other hand, when untrusted App 2 presents the same request (3), the hypervisor returns instances of A’ and C (4). Since C and other resources D, E, F and G are not related to the security goal, they are not virtualized, and either app may share them. Thus, the isolation size is minimized to just A according to the threat, and resources that can and must be shared for transparent operation remain as shared as intended. Moreover, Framework complexities that would arise from utilizing kernel-level isolation mechanisms are completely avoided.

A non-trivial challenge that can sometimes arise when PINPOINTing certain resources is when there are other operating system components or resources, shown in Figure 1 as resource G, that depend on interactions (5) with A or B and are unaware that now multiple virtual copies of them exist. In these cases, G must also be modified to account for this. Since this represents additional complexity, one must always consider whether this extra complexity will negate the lightweight benefits of the PINPOINT approach. If A is a large and complex object that has many dependencies throughout the system, it’s likely that creating virtual copies of A will break many things that assume there is only one A. In cases like these, it may be better to use a coarser isolation such as those discussed in Section VI. On the other hand, if A has few dependencies, then the modifications to G (if in fact there are any) will be straightforward. Two of our four case study applications described in Section IV exhibit this characteristic, and these details will be included there.

Another challenge that can arise is when the same sensitive

---

1We use the term hypervisor to indicate the relatively small scope of authority compared with supervisors (i.e., kernels) that have authority over an entire userspace, and type I (native) hypervisors that have authority over one or more guest operating systems. The term has been used similarly in [23].
information can be revealed by more than one object. For example, let’s say both A and C are capable of returning a piece of sensitive data such as IMEI. It is important that all of these paths be identified, and either blocked or added to the isolation boundary. Our case study encountered one example of this which will be described in Section IV.

When designing and implementing the hypovisor, care must be taken to ensure that the system does not allow any form of delegation of the hypovisor’s duties. For example, if the hypovisor is responsible for dispatching a capability, there must be no other ways for an entity to acquire that capability. Any other ways must be blocked in order to maintain the integrity of the isolation. Section III-C contains a specific example of this and how we addressed it using mandatory access controls (MAC).

Identifying the best place to instantiate the hypovisor is key to achieving a balance between flexibility and specificity. In our experiences thus far, we have found that the best isolation points are places in the Framework where classes of objects and/or their capabilities are managed or dispatched to apps. In Android, many resources are abstracted as system services, and their capabilities are dispatched by Context Manager (i.e., the native servicemanager process). As such, our initial work has focused on PINPOINTing system services, and our accomplishments thus far in this regard are described in Sections III and IV. However, we see future opportunities for implementing complementary hypovisors in key places other than system services, including:

1) High level data objects, such as ContentProvider. These may leak personal data [25].
2) Binder and Intents. These may be used as a path for a malicious app to attack or trick other apps [15].
3) Camera, audio. These have obvious privacy implications if misused, e.g., [30].
4) Clipboard. Can be used as an attack channel [35].
5) Accessibility subsystem. May be malicious toward critical apps [20].
6) Notifications. Potential misuse [34].

III. CASE STUDY ON ANDROID SYSTEM SERVICE

In order to evaluate the PINPOINT concept, we undertook a case study using Android’s system services framework in both Android 4.4.4 (KitKat) and 5.1 (Lollipop) on a Nexus 5 device. Since a wide variety of key resources are abstracted as system services, this choice illustrates that if the point of isolation is wisely chosen, a single PINPOINT hypovisor can be used for a variety of situations. To illustrate this point, we first provide some background on system services.

A. Android System Services

Interactions between Android applications and system services is enabled by the Binder and Service Manager subsystems. Binder relies on capability-based security and implements a “call by invitation” mechanism to allow communication among apps, system services and Service Manager. As such, before an app is allowed to call a service, it must receive an invitation in the form of an IBinder token.

Invitations are first created when services are registered with the central directory of services known as Context Manager. By design, there can be only one Context Manager, a designation granted exclusively to the native servicemanager process early during the boot process, by way of its privileged relationship with Binder. Once Service Manager becomes Context Manager, System Server registers core system services using the addService() method of Service Manager. The result of this registration process is that Service Manager now holds an invitation (IBinder) for every system service running on the device. When an app needs an invitation for one of these services, it contacts Context Manager. Context Manager then passes this invitation to the app, and upon seeing this transaction, Binder updates its protected list of invitations held by the app. Invitations cannot be forged because any forged invitation will not have a corresponding entry in the protected list maintained by Binder.

All requests for system services, even those made by system components, must go through Context Manager. Thanks to Binder, the native servicemanager process has access to the trusted identity of the caller, in the form of the Linux uid, which corresponds to the Android userId and appId. This makes servicemanager an excellent place to implement a system service hypovisor that can regulate applications’ interactions with virtualized system services. In this way, this hypovisor represents the PINPOINT “sweet spot” of being specific enough to limit inter-dependencies with other parts of the system, but flexible enough to apply to a large class of objects and the mechanism whereby their capabilities are dispatched.

B. PINPOINTing System Services

We present our design in three parts as shown in Figure 2, beginning with the central enabling core, the system service hypovisor, labeled ①. The hypovisor exists within the native servicemanager process, where all service lookups are processed and capabilities dispatched. Lookup requests by

---

TABLE I. SUMMARY OF NAMESPACE TRAITS AND THEIR VALUE TO ANDROID SECURITY

<table>
<thead>
<tr>
<th>Namespace Trait</th>
<th>Value to Android Security</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine-grained isolation of specific resources</td>
<td>Tailored isolation environment for each application, addressing specific threat and/or user goal</td>
</tr>
<tr>
<td>Resource-centric isolation</td>
<td>Match user perspective on security; increase usability; simplicity</td>
</tr>
<tr>
<td>High efficiency</td>
<td>Negligible performance impact; design simplicity</td>
</tr>
<tr>
<td>Share-by-default</td>
<td>Preserve open system design; avoid breaking things unrelated to the isolated resource</td>
</tr>
<tr>
<td>Transparent to host and apps</td>
<td>System retains control over apps; apps run unmodified</td>
</tr>
<tr>
<td>Small footprint (files, memory)</td>
<td>Little impact on performance &amp; resources; OTA updates</td>
</tr>
</tbody>
</table>

---

2In fact, in the Android source (frameworks/native/cmds/servicemanager/binder.h), Service Manager is described as “The One Magic Object”.

---

3
Every process using Binder, including system service threads within system_server has a protected representation in the kernel as an instance of a binder_proc structure. Each remote capability that a process holds is represented by one or more binder_node structures attached to the binder_proc instance. These nodes are known only to the kernel module and are used to determine the recipient of the communication, based on a handle provided from userspace. Handles are local references and mappings from handles to nodes are also stored securely in binder_proc. Hence, only the kernel knows how to map a particular handle to the corresponding node.

When System Server registers system services with Context Manager using addService(), the kernel adds the service’s binder_node to servicemanager’s binder_proc. Context Manager also allocates a local index to each registered service. When an app asks Context Manager for a handle to a service, servicemanager returns the handle and the kernel binder driver adds the service’s binder_node to the app’s binder_proc. Apps can also send handles they possess to other apps via Intent. Upon seeing the handle within the transaction, the kernel driver adds the node to the receiver’s binder_proc so that the recipient is now a valid holder of that capability. This is known as a binder transfer.

With our addition of a services hypervisor, we do not change anything about how handles are looked up and provided by Context Manager or how capabilities are propagated by way of the kernel binder driver. All apps, native and Java alike, are subject to the intervention of our hypovisor when requesting service handles. Any vulnerabilities in our prototype regarding how service handles are obtained by apps, or vulnerabilities in the binder driver itself, are also vulnerabilities of stock Android and thus outside the scope of this discussion.

What our design does change is which handles are given out. In PINPOINTing, services, we have introduced the notion of remote service handles that should be unobtainable by certain apps. This is different than stock Android where Context Manager acts as an open directory service, and obtaining a service handle via binder transfer from another app does not represent a capability leak. In our design, this rather unusual case of app-to-app transfers of system service handles must be prevented so that our hypovisor cannot be bypassed. Our prototype achieves this blocking in the kernel binder driver’s binder_transaction() function, through an extension of existing SEAndroid MAC. Specifically, we extend the security_binder_transferمخاط() hook by also passing the task_struct of the binder_ref (for references) or binder_node (for handles) being transferred, so that the hook function can extract the owner’s security identifier (SID) and decide if the transfer should be allowed. Finally, we modified the type enforcement rules associated with untrusted_app to disallow transfer of u:system_server:s0 binders between u:untrusted_app:s0 apps. By adding a neverallow rule, we further ensure at build-time that there are no allow rules elsewhere in the policy that are inconsistent with this. This effectively blocks any attempted bypass of our hypovisor, while allowing all other normal binder transfers among apps and the system to proceed.

C. Security Discussion

In our PINPOINT case study, we introduce a lightweight hypervisor into the native portion of Service Manager. The purpose of the hypovisor is to isolate particular apps from various services as specified by the user’s policy. This security discussion is included to provide a sense of the strength of this isolation. We begin with Binder, since most of the isolation strength derives from Binder’s security model.
D. Policy Configuration

As explained in Section III-B, the hypervisor within `servicemanager` consults a secure policy file to determine if the requester has been assigned to any virtual services. This policy can be created and updated by a variety of means: via the system Settings app, via launcher configuration, from hard-coded (i.e., build-time) mandatory policy, via over-the-air (OTA) updates in a mobile device management (MDM) architecture, etc. In our prototype, we included a default policy file in the system build, and updated it via `adb` and a custom launcher application. In terms of user-friendliness, our custom launcher enables the end-user to drag-and-drop app icons to and from different containers, each representing a specific PINPOINT configuration. For example, a particular container might be configured to protect two sensitive resources, location and IMEI, from the apps placed within it. When an app is dropped into this container, the launcher app automatically updates the policy with the `uid` and service names corresponding to the protected resources. This update takes effect immediately since `servicemanager` consults the policy each time the app makes a request.

E. Limitations

Currently, our case study prototype requires all global and virtual system services to running whether or not any apps are assigned to them. In terms of overhead, this fact manifests itself as additional memory use by the `system_server` process. Although we show in Section V-A that this overhead is small, we feel that this aspect of the design could be made more elegant and efficient in the future.

It is also important to note that our design does not provide full security domain isolation in the sense that it does not prevent apps from passing high-level sensitive information to other apps.

IV. Applications

In this section, we describe the specifics of our experience with PINPOINTing four common system services, based on specific security goals. All implementations were tested in Android 4.4.4 (KitKat) and then ported to Android 5.1 (Lollipop) on a Nexus 5 device.

1) `LocationManagerService`: A widely used location-finding service that binds with a number of abstract provider mechanisms. Security goal is to prevent untrusted apps from obtaining accurate location information [8]. See Section IV-A.

2) `iPhoneSubInfo`:: A “hidden” service for accessing phone subscriber information, called only by other system services such as `TelephonyManager`. Security goal is to prevent untrusted apps from accessing sensitive subscriber information [12]. See Section IV-B.

3) `InputMethodManagerService`: A service that arbitrates communications between apps and a variety of installed input methods, and has complex interactions with other system objects including `WindowManager`. Security goal is to protect critical apps from falling victim to malicious input methods [27]. See Section IV-C.

4) `SensorService`:: A native service that interfaces directly with hardware devices. Security goal is to prevent untrusted apps from obtaining accurate sensor data to steal data [33] [3], eavesdrop [24], or track movement/location [22]. See Section IV-D.

By and large, porting from 4.4.4 to 5.1 was straightforward. In one case however (`iPhoneSubInfo`), changes to the underlying service architecture required us to slightly redesign and expand the isolation boundary in order to continue to meet the security goal. This will be discussed below.

A. Location Service

Although location services provide great convenience and enable new functionality for users, they have significant security and privacy implications if misused. While some apps require accurate location to fulfill their main purpose, others utilize location information only to enrich their primary function. For example, a social networking app’s primary function is to interact with friends via photo and status updates. These apps usually enrich this interaction by attaching location to these updates. If the end-user wishes to prevent only this one app from knowing location, and still enjoy its primary friend-interaction functions, she must rely on the trustworthiness of the app’s own settings and controls. This is because current location privacy support from Android itself is too coarse-grained to achieve the user’s goal of isolating only this one aspect of this one app. If the app is poorly-written or malicious in its handling of location data, privacy leaks may occur despite the user’s best efforts to prevent them. By PINPOINTing the location service resource, and placing only this app in the new location namespace, we can transparently and effectively address this user’s security goal without inconveniencing her or introducing the complex system modifications and overhead of general-purpose solutions.

To demonstrate this, we PINPOINTed the location service to provide three separate location namespaces for assigning apps, each with different semantics but identical interfaces. The global location namespace functions normally and is used with trusted apps. A fuzzy location namespace provides reduced-accuracy location information by adding noise to location objects. Finally, a random location namespace returns totally random location data to assigned apps.

We implemented these two additional location namespaces by adding two additional system services, `LocationManagerService_1` (LMS’), and `LocationManagerService_2` (LMS”), as shown in Figure 3. These present the exact same API as the stock service, and thus are indistinguishable from the app’s perspective.

Each location service binds to the standard set of common location providers such as `GpsLocationProvider` that interfaces through native code to actual hardware. However, as alluded to in Section II, these providers represent dependent resources (`G` in Figure 1) that are designed based on an assumption of only one location service. Thus, these must also be modified slightly to make callbacks to all three location services. Otherwise, LMS’ and LMS” will never get location update callbacks since the providers are not otherwise aware of the virtualized services.
Fig. 3. PINPOINTing LocationManagerService.

(a) RunKeeper running in location namespace with added noise.
(b) RunKeeper running in random location namespace.

The semantics of the additional services are as follows: LocationManagerService_1 replaces location updates from the providers with random data, while LocationManagerService_2 adds random offsets to the same. Since each namespace is indistinguishable from the global location namespace in, apps in alternate namespaces behave normally and process the virtual location data as if it were real.

Figure 4 shows screenshots of a popular fitness app, RunKeeper, that we used to demonstrate the isolated location namespaces. Figure 4(a) shows points collected during an activity while the app’s uid is assigned the noisy location namespace. Figure 4(b) shows the same app while assigned to the random location namespace. Note that in both cases, the app’s display indicates “Good GPS”, demonstrating the complete transparency of these namespaces to this unmodified app.

B. Subscriber Information Service

iphonesubinfo is a hidden service used exclusively by TelephonyManager to service app requests for subscriber information such as IMEI, mobile equipment identifier (MEID), electronic serial number (ESN), phone number, voicemail number, private/public user identities, home network name, etc. Several of these values have significant security and privacy implications and are known to be malware targets [12]. Although protected by Android’s READ_PHONE_STATE permission, misusing or malicious apps can easily legitimize declaration of this permission since it is necessary for a number of common features, such as those provided by PhoneStateListener.

To isolate an untrusted app from or more of the data values returned by iphonesubinfo, we PINPOINTed this system service. We enabled the non-global namespace by modifying the internal telephony ProxyController to instantiate PhoneSubInfoController_1 as well as PhoneSubInfoController. The former starts iphonesubinfo_1 service with an API identical to iphonesubinfo, started by the latter. When an untrusted app is assigned to the alternate iphonesubinfo namespace, it can obtain the same instance of TelephonyManager as trusted apps can, but any subsequent calls to getDeviceId(), getLine1Number(), etc. by the untrusted app are processed by iphonesubinfo_1. iphonesubinfo_1 returns different values for sensitive subscriber parameters.

When porting this design to Android 5.1, we found that the underlying structure of the telephony service had changed significantly. In particular, the ITelephony (phone service) interface was enhanced to include its own getDeviceId() call, and TelephonyManager was modified to obtain the device ID from this interface rather than IPhoneSubInfo as was the case in 4.4.4. Thus, apps assigned to iphonesubinfo_1 would still get the device’s real IMEI because our isolation did not include every object that could return that sensitive data. This necessitates an expansion of the isolation boundary to include both iphonesubinfo and phone services, and is a good example of needing to identify all possible means of access to the sensitive resource related to the end security goal.

To demonstrate effectiveness of our PINPOINTed subscriber information service, we obtained the popular app Your SIM Card. Figure 5 shows this app running unmodified in both global (Figure 5(a)) and fake (Figure 5(b)) iphonesubinfo/phone namespaces. In the global namespace, the actual, valid IMEI of our test device is returned, while a fake IMEI is returned to the app after it has been assigned to the alternate iphonesubinfo_1/phone namespace by adding its uid to the nsPolicy file.

C. Input Method Service

Input Method Editors (IME) are screen controls that enable users to enter text. Currently, there are about 900 third-party keyboard apps published on the Google Play store, with at least 10 having more than one million downloads. Most require INTERNET or WRITE_EXTERNAL_STORAGE permissions, which enable the IME to log or transmit any data that’s typed in. In an empirical study of keyboard apps, it was found that more than 61% require three or more permissions giving them the ability to exploit keylogging and man-in-the-middle attack
The overall working architecture of IMEs is shown in Figure 6. In every application’s context space, there exists an instance of InputMethodManager (path 1) which is used to communicate with a system-wide service, InputMethodManagerService. When an input field comes into focus, the app’s InputMethodManager invokes this system service (paths 4 and 5) after obtaining its handle via Service Manager (paths 2 and 3). With this handle, the app may obtain a unique InputConnection Binder token from InputMethodManagerService for making direct calls to the IME keyboard app. Using this token, the system is able to secure and control interactions among multiple applications and multiple IMEs [18].

Currently, apps do not have control over the IME selected by the user. Instead, the system will bring up the user’s selected IME whenever any text field comes into focus. While Google has recognized the security and privacy issues associated with this design [1], the current measures rely on the user to make wise choices regarding IME installation and selection. Using session information attached to each window instance by WindowManager, the Input Method Framework (IMF) ensures that only the active activity can get access to the data being entered. Furthermore, InputMethodManagerService ensures that all messages received from running IME applications are from the current user. Importantly, this includes messages for changing IMEs (i.e., messages resulting from calls to InputMethodManager.setInputMethod()), which are guarded with the token to ensure that they originated from explicit user selection. However, none of these protections will help if the IME itself is malicious or compromised and the user selects it.

By PINPOINTing the InputMethodManagerService, we are able to provide a mechanism to shield sensitive apps from falling victim to a malicious IME selected by a tricked user. Figure 7 shows the PINPOINT concept applied to input methods. This is accomplished by using the our PINPOINT service hypervisor prototype to virtualize InputMethodManagerService. In the figure, IMMS corresponds to the “real” InputMethodManagerService (input_service), while IMMS’ is a second service (input_service_1), with an identical interface and features except for the fact that it holds only a subset of all available IMEs.

As suggested in Section II, there are additional complexities with virtualizing IMEs due to dependencies with other objects in the system. Because of interactions with WindowManager mentioned above, we needed to make minor modifications to WindowManager, so that it can be aware of all the InputMethodManagerService namespaces running and push updates about the current activity to all of them. As with location service, this situation corresponds to dependent resource G in Figure 1. To enable independent InputConnection from each app’s InputMethodManager instance to each service, we created a Java interface which all of the InputMethodManagerService instances implemented.

A demonstration of our IME namespaces is shown in Figure 8. Figure 8(a) depicts a non-critical app, EatStreet, assigned to the global IME namespace, where any IME can be used, including a representative untrusted IME, SwiftKey (added from Google Play). Here, the Choose input method dialog shows all installed input methods. In contrast, the critical banking app of Figure 8(b) has been assigned to the alternate IME namespace in order to protect its data from possible malicious IMEs. Hence, the chooser only allows selection of trusted IMEs, while SwiftKey is excluded as an authorized IME for this app.
D. Sensor Service

Modern mobile devices have a rich set of environmental and motion sensors available to apps. Unfortunately, the Android security architecture does not extend to most of these sensors, making it all too easy for malware to utilize them to compromise user data entry [33] [3], eavesdrop on voice communications [24], track user movements, and infer location [22]. By PINPOINTing SensorService, we enable the user to take advantage of apps without needing to also trust their handling of sensor data.

In the Android platform, apps may acquire sensor data by getting an instance of SensorManager, which in turn accesses raw sensor data via SensorService, a native system service. SensorService’s threadLoop() collects raw sensor data in a structured data buffer of type sensor_event_t, which is then returned to the app via its SensorManager’s SensorEventConnection. The buffer structure contains raw sensor data for each of the device’s sensors including acceleration, magnetic, orientation, gyro, temperature, distance, light, pressure, and relative humidity.

To PINPOINT sensor resources, we followed the same general approach as with previous examples, by adding two additional native SensorServices to the device, and registering them with Context Manager as sensorservice_1 and sensorservice_2. For demonstration purposes, we hardcoded sensorservice_1 to overwrite the gyro, magnetic, and orientation structure members of the buffer structure with random data before it is returned to the app’s SensorManager. Likewise, sensorservice_2 is hardcoded to overwrite only the light structure member of the structure with random values. Structure members containing data from other sensors are passed through unmodified.

With three possible sensor service handles on the device, SensorManagers of apps assigned to one of the two alternate sensor namespaces are always given handles to sensorservice_1 or sensorservice_2, depending on their assignment. To demonstrate the effectiveness of this, we downloaded AndroSensor a popular Google Play Store app, and ran it in each of the three sensor namespaces. Figure 9 shows AndroSensor running in the global sensor namespace, with all sensor traces steady, indicating a stable physical environment. In contrast, Figures 10(a) and 10(b) show AndroSensor running in the alternate sensor namespaces of sensorservice_1 and sensorservice_2, respectively. For all three cases, the physical environment was approximately the same.

V. Evaluation

A. Performance

To evaluate the overall performance impact of PINPOINTing services, we performed the benchmark tests shown in Table II, with and without namespaces. For each benchmark, we measured performance under four different device configurations: 0NS represents stock Android without any PINPOINT capability or namespaces, while 1NS, 2NS, and 3NS represent devices configured with one, two and three PINPOINTed services, respectively. Figure 11 shows the average value of 10 runs of each benchmarking test.
TABLE II. EVALUATION BENCHMARKS USED.

<table>
<thead>
<tr>
<th>Name</th>
<th>Version</th>
<th>Workload type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linpack</td>
<td>2.2.8</td>
<td>CPU</td>
</tr>
<tr>
<td>SunSpider</td>
<td>1.0.2</td>
<td>CPU &amp; I/O</td>
</tr>
<tr>
<td>Quadrant Advanced Edition</td>
<td>2.1.1</td>
<td>File I/O</td>
</tr>
<tr>
<td>Quadrant Advanced Edition</td>
<td>2.1.1</td>
<td>2D &amp; 3D</td>
</tr>
</tbody>
</table>

We also measured the impact on memory of adding PINPOINTed services. Since each running service represents additional threads within SystemServer, we measured VmSize of the system_server process by reading its /proc/<pid>/status under each of the same four configurations. Figure 12 shows the average value of 10 measurements of memory footprint for each configuration.

B. Discussion

Performance evaluation results presented in Section V indicate that increasing numbers of PINPOINTed services has no apparent effect on CPU, browser, or graphics performance. On the other hand, we observe a clear correlation between the number of PINPOINTed services and file I/O. Decreases in this score with increasing numbers of namespaces is expected due to an increase in policy file size and associated data structures being parsed and searched by servicemanager during every service lookup request in order to support namespace reassignments of running apps. In our current, unoptimized design, file I/O performance degrades by an average of 1.57% of the 0NS value for each additional namespace represented in the policy file. Although this degradation is negligible for simple policy files, we feel that this is an area for improvement. Our future implementations will include an optimization of this code, and policy options to configure how often policy lookups are performed.

We also observed a growth in system_server's memory size that is correlated to the number of additional service objects (i.e., namespaces) available for use in the system_server process. On average, we observed this increase to be approximately 0.64% of the 0NS value per each additional service. For a system with one additional IMEI namespace, two additional location namespaces, one additional input method namespace, and two additional sensor namespaces, system_server would have an approximately 3.84% larger memory footprint than the stock process. Note that an unused namespace still consumes additional memory, but since it does not add to the policy file, it will not contribute to file I/O degradation.

VI. RELATED WORK

A number of previous efforts have addressed the problem of untrusted apps having access to sensitive or private information. Some of these address specific types of data, such as location, while others look for more general solutions.

Two significant isolation approaches that influenced us tremendously are Cells [7] and AirBag [32]. Cells leverages Linux Namespaces to allow multiple Android user spaces to run simultaneously on a single hardware platform. Each user space, or virtual phone (VP), is isolated in a combination of separate Linux Namespaces for file system paths, process identifiers, IPC identifiers, network interface names, user names, and hardware devices. Cells introduces the concept of a foreground and multiple background phones that are isolated from each other so that malicious or buggy apps in one VP cannot affect others. Isolation in Cells is thus achieved at the virtual phone boundary.

AirBag also leverages Linux Namespaces, but achieves isolation at the native runtime boundary. This is accomplished by instantiating a separate app runtime that has virtually no interaction with the original native runtime. Each isolated runtime contains its own copies of key service processes and daemons, such as vold, binder and servicemanager that are launched in separate namespaces as compared with the normal runtime. Thus, an untrusted app “sees” an entirely different set of services, binder objects, file paths, etc. through the lens of its decoupled runtime. The untrusted app cannot communicate with apps in different runtimes, and the system resources it can view and control are completely dictated by the isolated runtime. Condroid [4] improves on AirBag’s design by restoring binder communications via virtual binders and increasing efficiency by enabling many system services to be shared among runtimes instead of duplicated.

While Cells, Airbag and Condroid provide excellent general-purpose isolation, their designs are complex, burden the system, and somewhat intrusive from the user’s point of view. For example, all three require special modifications to numerous shared hardware drivers, duplication of system processes and resources not related to the security goals, and introduce significant usability restrictions. Although they leverage lightweight Linux Namespace isolation, key benefits of Namespaces (Table I) are lost when the Android Framework is added on top since many fundamental aspects of Android’s
open design are broken by the kernel-level isolation. Fixing these problems greatly complicates the designs. Thus, our main difference from these works is our deliberate choice not to provide a general-purpose solution, but rather one that addresses specific security goals by directly isolating the specific Framework objects associated with the security goals. For these specific cases, our approach is simpler, less burdensome, and more usable.

MOSES [29] is a framework designed to isolate applications and data for the purpose of protecting sensitive corporate data. While MOSES also represents an effective general solution to securing corporate data leaks on mixed-use personal/business devices, it is not very suitable for protecting users’ privacy or securing specific resources because of its security profile-centric architecture that forces explicit switching and carries performance penalties.

IPC Inspection [15] and Quire [9] prevent privilege escalation and confused deputy attacks among apps and do not address an app’s direct access to resources it already has adequate permissions for, as our work does. TaintDroid [11] inspects and analyzes information flows across the system, but does not provide the means to manipulate or block information. AppFence [19] leverages TaintDroid monitoring to enable data substitution and blocking. For information resources, such as location and IMEI, the resulting capability is similar to some of our basic namespaces. However, AppFence cannot control the semantics of functional resources, such as we demonstrated with the input method namespace. Furthermore, AppFence’s substitution and blocking capabilities affect information resources for the entire platform rather than being selective for individual apps as is the case in our system services case study.

Mr. Hide [21] adds finer-grained permissions to apps by way of byte code rewriting, while APEX [28] introduces context-sensitive run-time permissions. Compac [31] allows different components within apps, such as in-app ads, to have different sets of permissions. These and other permission-enhancement designs can only restrict access to resources and are unable to redefine them as we do.

Finally, as location data is widely viewed as having serious privacy implications, there are numerous works specific to improving location privacy. LP-Guardian [14], LISA [5], and Koi [17] are examples of these. While each is effective for controlling or preventing the use of location data, they are not generally applicable to other resources as PINPOINT is. As our case study on PINPOINTing system services demonstrates, if the point of virtualization is chosen wisely, the resulting isolation capability is flexible enough to apply to classes of resources rather than only specific ones as these works do.

VII. Future Directions

In our present work, we have gained a tremendous insight into the trade-off between isolation design alternatives, system complexity, usability, convenience and effectiveness. We plan to further quantify these relationships so that we can make informed choices when addressing the high-level requirements typically stated by end-users. Ultimately, we plan to formalize the PINPOINT methodology, so that security designers can easily understand the trade space of PINPOINT designs vs. general-purpose approaches.

Through our case study of implementing a services hypervisor, we’ve acquired a sense for the difficulty of implementing a representative PINPOINT hypervisor and its companion virtual resources within the Android Framework. Encouraged by our experiences, we plan to consider the potential benefits of PINPOINTing other resources, including those outside the purview of Service Manager. Following from this, we envision implementing a container abstraction, whereby multiple, heterogeneous PINPOINTs, Linux Namespaces, and other forms of access control and virtualization can be easily combined by the end-user to form easily-understood security and privacy macros such as “incognito,” “banking,” etc.

We recognize the inflexibility of having to define PINPOINTed resources at system build time. As such, we see opportunities to investigate techniques for establishing new PINPOINTs while the system is running. Also, we would like to evolve our current rudimentary means of policy configuration into a more powerful and intuitive means for end-users to configure, combine and use PINPOINTed resources, possibly through an advanced launcher interface.

Additional project details, status, and instructions for requesting access to our prototype code are available at https://goo.gl/2pJeMp.

VIII. Conclusion

We have presented PINPOINT, a Android resource isolation strategy that forgoes general-purpose solutions in favor of a “building block” approach, similar in concept to Linux Namespaces, but implemented in the Android Framework. By addressing stated security goals and no more, PINPOINT yields an effective result using only the minimum amount of isolation. This helps minimize or eliminate the negative side-effects that are sure to emerge when large parts of Android’s open architecture are subject to isolation.

Through our case study on Android System Services, we uncover the primary considerations of a PINPOINT designer. These include correlating the stated security goals with specific Android resources, identifying all Framework objects that have access to these resources, and finding any dependent resources that may exist in the remainder of the system. With this insight, an appropriate hypervisor is then implemented, corresponding objects with alternate semantics are created, and system MACs are updated as necessary to enforce the hypervisor’s authority.

Our system services prototype and experiments with location, subscriber, input method and sensor services demonstrate that for specific security or privacy goals, PINPOINT yields effective solutions for unmodified apps with a minimal amount of design complexity, system modifications, or negative impacts on user experience.

Acknowledgement

We thank the anonymous reviewers for their careful reading of our manuscript and the many insightful comments and suggestions. This work was supported in part by NSF Grant 1318814, and AFRL project GAICHCYBR. Any opinions, findings, conclusions or recommendations expressed in this material are solely those of the authors and do not necessarily reflect the views of the NSF or the US Air Force.
REFERENCES


