The Lookout Security Cloud

Cloud-first, device-assisted mobile threat defense
The Lookout Security Cloud

I. Introduction

The Lookout Security Cloud is a cloud-based platform that detects and stops both broad-based and advanced mobile threats. It protects mobile endpoints and infrastructures from threats across the mobile risk spectrum, enables deep threat investigation, and ultimately powers a wide range of Lookout products:

The Lookout Product Family Architecture

What makes the Lookout Security Cloud different? Four key areas differentiate Lookout’s platform from other vendors in the mobile threat defense space:

1. Our Massive Global Device Network

The success of Lookout’s personal and enterprise endpoint products has allowed Lookout to gather security data from over 100 million mobile devices worldwide. Every month tens of millions of devices in over 150 countries send security telemetry to the Lookout Security Cloud, ensuring that Lookout continues to lead the industry in novel threat discoveries like the Pegasus mobile spyware.

2. Our Industry-Leading Mobile Dataset & Machine Intelligence

Machine intelligence only matters insofar as you have a large enough global dataset to train machine threat models to understand if a potential threat signal or characteristic is normal, rare, or truly anomalous in the world. Lookout leads the mobile threat defense industry in this regard, with over 111 granted patents related to mobile security and the industry’s largest dataset, having acquired and analyzed, for example, over 40 million unique mobile applications in the world.

3. Our Comprehensive Security Capabilities

One platform, the Lookout Security Cloud, powers a wide range of security products and services and delivers comprehensive capabilities for mobile threat defense, mobile app reputation services, and mobile threat intelligence.

4. Our Cloud-First, Device-Assisted Security Approach

Mobile-first since 2007, Lookout knows how to build security optimized for mobile environments. Since on-device analysis techniques are inherently less secure (because attackers can reverse engineer them) and they can also negatively impact device performance if over relied upon, Lookout built a cloud-first threat analysis engine to complement our on-device security technologies, creating a faster, more secure, and ultimately more performant mobile security approach.

The next sections of this whitepaper will examine some of Lookout’s specific security technologies that keep both mobile enterprises and individuals secure.
II. App Analysis Architecture

The diagram below depicts the architecture of the platform’s app-based threat detection capabilities. This architecture follows a four-step process: (1) data acquisition (2) data enrichment (3) data analysis (4) protection.
1. Acquisition

The platform collects real-time security telemetry on mobile applications from a variety of sources:

**Mobile Sensor Network**  More than 100 million registered mobile devices worldwide have provided Lookout with a comprehensive, real-time view into threats on just one device or millions. Lookout’s app binary acquisition process spreads the load among multiple devices to limit battery and data impact, reassembling the app fragments in the cloud and preserving end-user privacy by only collecting application binaries, not user personal data (e.g. photos, messages) generated in the course of using these applications.

**Crawling**  Lookout continually monitors the major and minor app stores of the world, including app stores in countries such as China, Russia, and India. Lookout’s crawling technology also enables app acquisition from ad hoc web sources.

**APIs**  By serving as the exclusive security layer for some of the world’s largest app stores, the Lookout Security Cloud has privileged access to malware submitted to these stores that never sees the light of day.

---

Lookout’s platform is aware of the presence of 67,500,000 unique app binaries in the world, counted by cryptographic hash. This include both system apps (apps that are part of the operating system) as well as user-downloaded apps, and counts each version of an app as a unique app instance.

---

<table>
<thead>
<tr>
<th>At a Glance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Devices that have contributed security telemetry to Lookout</td>
</tr>
<tr>
<td>API Partners</td>
</tr>
<tr>
<td>Unique app binaries detected</td>
</tr>
<tr>
<td>Unique app binaries acquired</td>
</tr>
<tr>
<td>Unique app binaries detected on only one device worldwide</td>
</tr>
<tr>
<td>Apps acquired daily</td>
</tr>
</tbody>
</table>
2. Enrichment

Each app acquired by Lookout’s platform undergoes a unique enrichment process that characterizes how it works and accurately relates it to the world of known applications:

**Metadata** Lookout appends data that includes app name, digital signature, app store description, and developer name.

### METADATA EXAMPLES

- **Package name:** com.android.service
- **Signer:** bb626d3b8406e7fc330d0f4b304cbfc5f610721f
- **Signer metadata:** CN=Dragon, L=SZ, ST=GZ, C=CN
- **Packaged date:** 2012-09-20 18:36:44 UTC
- **Signed date:** 2012-09-20 18:36:42 UTC

**Reputation** Lookout incorporates data related to the authorship, origin, and geo-historical distribution of an app, such as the duration and location of its popularity.

### REPUTATION EXAMPLES

**REPUTATION RESULTS:**

- 95% of known APKs that use this signer are malware.

**Behavior Analysis** The platform generates app behavior data, generated through dynamic and symbolic execution technologies that run the app in a simulated environment and analyze the capabilities of its code.

### BEHAVIOR EXAMPLES

**BEHAVIOR ANALYSIS RESULTS:**

- write_file (Osiris[0.1.217])
- read_contacts (Static Behavior Extraction[3.1.469])
- write_contacts (Static Behavior Extraction[3.1.469])
- read_sms (Static Behavior Extraction[3.1.469])
- read_imsi (Static Behavior Extraction[3.1.469])

**Binary Similarity Analysis** The platform automatically assesses the fuzzy code similarity an app shares with all known code in Lookout’s mobile intelligence dataset. It reveals where that app’s code (or its relatives) appear in the world by analyzing approximate similarity between individual code classes and then computing an aggregate similarity score.

### BINARY SIMILARITY EXAMPLES

<table>
<thead>
<tr>
<th>INDEX CLASS</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lorg/linphone/MapAPP$1$</td>
<td>0.9433</td>
</tr>
<tr>
<td>Lorg/linphone/MapAPP</td>
<td>0.9846</td>
</tr>
<tr>
<td>Lorg/linphone/util/Constant</td>
<td>1.0000</td>
</tr>
<tr>
<td>Index match</td>
<td>0.9923</td>
</tr>
</tbody>
</table>
Lookout holds patents related to its binary similarity analysis technology, which is one of the key differentiating technologies that powers Lookout’s machine-intelligence-driven security model. Whereas attackers can evade signatures by changing a single line of code, Lookout’s binary similarity technology (known as App Genome Sequencing) does not depend on precise 1:1 matches and can assess approximate matches at both a granular (class or code block) and holistic (app) level. This dramatically raises the cost of attack because it requires attackers to essentially overhaul their entire code base to evade detection.

Even some of the less powerful enrichment technologies can play a key role in identifying and tracking malicious code by adding relevant data points to feed Lookout’s security engine and enable it to find more complex, multidimensional correlations.

3. Analysis

Lookout’s security engine ingests the data generated by the platform’s acquisition and enrichment processes and then automatically compares these data points to the hundreds of millions of data points in Lookout’s mobile intelligence dataset.

Multidimensional threat correlation makes the platform substantially harder to evade because it requires attackers to re-implement their entire platform and command and control infrastructure, instead of simply changing the few components that match a signature or obscuring the malicious activity that would trigger an alert. In the event that the Lookout Security Cloud finds no correlations the platform relies on a risk-scoring model, taking inputs from the enrichment and analysis processes to predict zero-day threats.

The stunning breadth and complexity of the multidimensional correlations generated by Lookout’s machine-intelligence-driven security far outpace the capacities of human analysts and behavioral analysis models alone.

Consider the diagrams on the following pages that visualize these correlations for two distinct malware families, Mouabab and NotInstalledYo.
Multidimensional Threat Correlation Analysis of Moubad Malware Family:

This diagram shows samples of the Moubad mobile malware family, correlated by shared signer, IP communications, and binary similarity as calculated by the platform's App Genome Sequencing technology. Moubad is a family of trojans that enable third party control over a compromised device, allowing remote attackers to send premium rate SMS messages and engage in remote dialing activities.
Multidimensional Threat Correlation Analysis of NotInstalledYo Malware Family:

This diagram shows samples of the NotInstalledYo mobile malware family, correlated by shared signers and binary similarity as calculated by the platform's App Genome Sequencing technology. The node at the center of this galaxy represents a widely shared signer that uses a compromised signing key. NotInstalledYo is a family of spyware that intercepts SMS messages on victimized devices and forwards them to attackers.

Red Zone Area Enlarged:

Samples that share a high degree of binary similarity are grouped by color and nodes to which multiple colored nodes connect signify a shared signer amongst those samples.
4. Protection

The output of Lookout’s platform is a dynamic security decision that identifies evolving known threats as well as unique, targeted attacks. When the platform detects novel threats it automatically initiates an investigative process, alerting Lookout’s Research and Response team to further investigate the operation and motivation of attackers, take remedial action such as issue server takedown requests, and ensure that relevant partners, customers and organizations take remedial action if needed.

This security telemetry includes:

1. **OS/Firmware data** - OS file metadata, such as the file name and hash
2. **Configuration data** - system properties of the OS configuration
3. **Device data** - device identifier information, for device remediation purposes

After collecting this data the platform then re-assembles it in the cloud to form a device fingerprint. It correlates the various data points of this fingerprint against Lookout’s mobile intelligence dataset to identify when a device is vulnerable or has been compromised, and can also predict device risk based on anomalies or correlations to known signals of compromise. When the platform detects a compromised device it executes remedial action through an integrated Mobile Device Management (MDM) client.

Today, most device compromise detection models rely on a handful of point tests, hard coded on the mobile client. Attackers have identified and successfully deconstructed these point tests and devised countermeasures to easily evade them. Lookout’s detection model, however, differs substantially from these approaches in that it collects a holistic fingerprint of the device profile and sends it up to the cloud to analyze on the server-side. Lookout’s security model offers two key advantages: instead of reverse-engineering a few client-side point tests, to evade Lookout, attackers need to mimic the entire device state and its corresponding signals, which significantly raises the cost of attack. In addition, the server-side analysis also inhibits attackers from easily reverse-engineering Lookout’s detection methodology.

IV. Device Analysis Architecture

The Lookout Security Cloud Device Analysis Architecture

To protect the underlying security of mobile devices from threats such as malicious rooting and jailbreaking, the Lookout Security Cloud collects a range of device security telemetry to form a digital fingerprint of each device.
V. Machine Intelligence in Action

The following case studies demonstrate how the Lookout Security Cloud has delivered on the promise of machine-intelligence-driven security and can detect threats for which no prior signatures exist and can even detect threats before they exhibit malicious behavior.

Case Study 1: BadNews

Consider the case of BadNews, a malicious mobile ad network. Lookout found BadNews embedded in 32 different apps that were live in Google Play and had received millions of downloads. BadNews enabled the installation of additional APKs and could open URLs in the browser, although it exhibited neither of these behaviors at the time of discovery. The Lookout Security Cloud, however, detected that BadNews contained code that shared statistically significant correlations to known Russian malware and, in a pre-crime maneuver, proactively protected Lookout-enabled devices.

Post protection, Lookout continued to monitor BadNews in the wild and later observed it distributing new zero-day trojans via the APK installation functionality. Notably, BadNews only engaged in this malicious activity for five minutes a day, effectively disguising its activity from sandboxed security environments where isolated, point-in-time behavioral analyses would not detect the activity. To read more about BadNews, please visit the Lookout blog.

Case Study 2: MalApp.D

The power of a predictive security model is evident in Lookout’s detection of MalApp.D, a mobile threat that matched no prior signature nor engaged in overtly malicious behavior, but nonetheless put enterprise contact data and voice communications at risk. MalApp.D was embedded in a seemingly benign VoIP app that was live in the Google Play Store at the time of Lookout’s detection. With a handful of positive reviews and a 4.2 star rating, the app appeared legitimate. Through multidimensional correlation, however, Lookout’s platform revealed that this VoIP app was likely developed by a known author of mobile malware and it therefore posed an unacceptable risk to enterprises given its access to device contacts and potential call recording capabilities. To read more about MalApp.D, please visit Lookout’s website.

VI. Conclusion

The Lookout Security Cloud analyzes potential mobile threats not in the context of a single server, a single device, or a single application, but in the context of global mobile devices and code.

The continued failure of signatures and behavioral analysis alone to consistently identify threats without oceans of false positives or false negatives reveals the critical importance of having large, contextual data sets. Lookout’s platform excels at finding the signal amid the noise because it has unprecedented insight into the code, both apps and firmware, running on tens of millions of devices around the planet. This massive dataset produces hundreds of millions of datapoints that the platform can use to correlate and predict security threats and risks.

Predictive security models require machine intelligence to identify exceedingly complex correlations and risk signals that humans cannot possibly identify at scale. Today, most detection systems excel only at identifying the bank robber who has already hit the vault. We should instead use the deluge of data available to us to predict the next bank robber based on their correlations across multiple dimensions to known bad actors.