Automating NIAP Requirements Testing for Mobile Apps

June 29, 2020

Homeland Security

Science and Technology

Automating National Information Assurance Partnership (NIAP) Requirements Testing for Mobile Apps

Authors: Department of Homeland Security Science and Technology Directorate
National Security Agency

June 29, 2020

Executive Summary

In the past decade, mobility has evolved from a differentiator or key enabler within the modern information technology (IT) enterprise to a business necessity and operational imperative. Organizations large and small, across all market sectors, have embraced mobility for its benefits, but in the process have assumed all of mobility's endemic risks as well. For federal agencies, the majority of which have made improved mobility core to their enterprise IT strategies, the stakes are particularly high given their critical role. In May 2017, the Department of Homeland Security (DHS) Science and Technology (S&T) Directorate, in consultation with the National Institute of Standards and Technology, published the Study on Mobile Device Security, a report to Congress that described the state of mobile device use within the federal government, the risks such use poses and how the risks might be mitigated. The report also identified remaining challenges. Among the risks described were vulnerable mobile applications (apps), which, the report notes, can be mitigated in several ways such as by developing apps compliant to validated security standards and using mobile app vetting tools and methodologies.

The DHS S&T-funded pilot, in partnership with NIAP, examined to what extent NIAP evaluations of mobile application software could be automated.

For many years, the National Security Agency (NSA)-funded National Information Assurance Partnership (NIAP) has been responsible for overseeing a program that certifies the security of commercial products used in National Security Systems (NSS). While NSS are a special category of systems whose requirements do not apply to most government IT, the success of NIAP's requirements and evaluation model has led many other agencies to adopt its standards as well as the results of its product evaluations when they make IT procurement decisions. Even so, some agencies may prefer a lightweight vetting process that enables them to quickly assess whether their myriad mobile apps comply with NIAP standards, while reserving full and thorough NIAP evaluation for their most critical and sensitive enterprise apps.

Government agencies, therefore, need policies and operating conditions that support two different kinds of approval approaches: those that support NIAP certifications that follow the official NIAP process and are managed under NIAP auspices (i.e., *NIAP certified*), versus approaches that are done independently of NIAP involvement, while using NIAP's standards and requirements (i.e., *NIAP compliant*).

Over the years, NIAP made several changes to its evaluation and certification processes with the goal of improving efficiency, cost-effectiveness, and throughput. Among these are greater flexibility in some aspects of the evaluation and certification processes as well as an increased emphasis on semi- or fully automated processes. More recent activities have focused on representing the security requirements in multiple formats for both human and machine consumption. These developments can prove beneficial not only for official NIAP certification; they can also benefit any government agency that wants to ensure that its apps are compliant to security best practices, including compliance with NIAP standards. This is because some agencies may not have the expertise, resources, or personnel to conduct analyses of apps against NIAP requirements, especially if they need to do this for many apps before approving their use. Some agencies already use automated tools from a variety of sources to help them perform app

vetting and will likely want to be able to continue to use automation to determine NIAP compliance.

DHS S&T, in partnership with NIAP, sponsored a pilot effort to determine to what extent NIAP evaluations (i.e., *NIAP certified)* of mobile app software could be automated. The pilot was funded as part of S&T's ongoing Mobile Security Research and Development program. During the pilot, Kryptowire LLC performed an automated analysis of Android and Apple iOS versions of the Intelligent Waves Virtual Mobile Infrastructure Platform Hypori application using their mobile app vetting infrastructure. Leidos, Inc. Common Criteria Testing Laboratory analyzed Kryptowire's results to determine whether they were consistent with the expected results of a conventional NIAP evaluation. NSA experts provided additional analysis of the findings.

The results are extremely promising as the pilot demonstrated that it is indeed possible to automate significant portions of the app software evaluation process, thereby increasing efficiencies, shortening approval times, and reducing costs. Figure 1 summarizes the pilot's findings and conclusions and demonstrates that the automated testing was, for the most part, able to accurately meet the intent of the NIAP requirements, with relatively small outlays of time, money, and personnel. Additional analysis by NSA experts concluded that most of the automated tests fully met the intent of the requirements (87 percent for iOS and 64 percent for Android). Others partially met the intent (20 percent for Android) of the requirement for a variety of reasons (e.g., did not gather enough data to unambiguously assess a pass or fail against), but could meet the intent with some implementation changes. A relative few did not meet the intent (13 percent for iOS and 4% for Android) at all (e.g., where a test produced the wrong kind of evidence or aspects of the requirement were ambiguous). Finally, some tests (12 percent for Android) were deemed "not applicable" (e.g., where the requirements were updated to remove or modify a test, but where these changes were not yet implemented in the Kryptowire product).

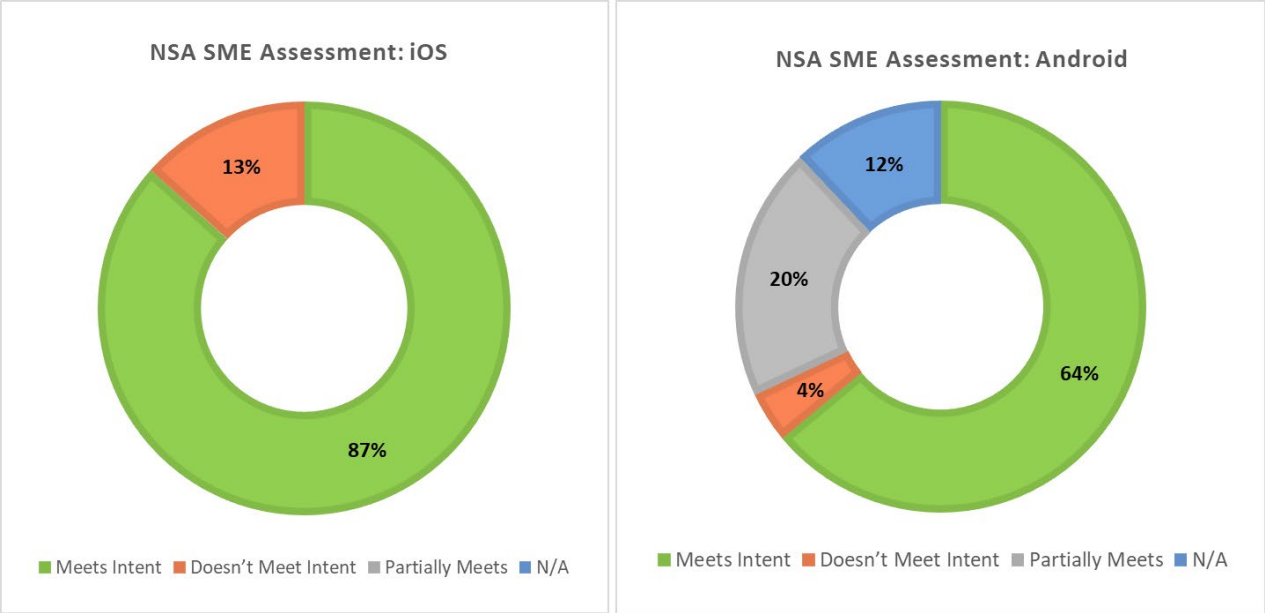

Figure 1. NSA SME Assessment of Whether Kryptowire's Evidence Meets Intent of Requirements.

The pilot also produced other findings regarding how NIAP certifications and NIAP-compliant app vetting can be designed and conducted in the future. These include:

- Automated app vetting against NIAP requirements enables successive updates to mobile apps to be tested and fielded faster. For NIAP certifications, this can be done without needing to undergo a complete NIAP evaluation each time and the updated automated test results could be included as part of the evidence and documentation provided to NIAP through the Assurance Maintenance process. For cases where the app is only assessed for NIAP compliance, agencies could examine the results and determine whether to approve the app.
- Automated vetting provides risk reduction for several stakeholders, including agencies, software vendors, and end-users wherein apps can be assessed ahead of time for basic compliance to the requirements before undergoing a formal NIAP evaluation.
- Apps can be accurately vetted even if analysts and evaluators do not have access to source code. Static and binary analysis can surface potential issues that are not obvious using other techniques.
- Agency approval authorities can benefit greatly from reduced risk even for commercial apps that will not undergo formal NIAP evaluation because they can identify and use apps that are compliant to NIAP standards and other best practices.
- Apps can be quickly vetted against any new or updated requirements, ensuring compliance to the latest NIAP standards/best practices and continued risk decisions, at speed.
- Some of NIAP's requirements and prescribed testing approaches, as defined in protection profiles, are not necessarily the best or most effective ways to test certain security requirements and there is a need for greater flexibility to exercise a variety of test procedures, while ensuring security.
- This successful collaboration among DHS S&T, NSA, and industry augurs well for other security automation efforts, some of which already are under way to improve automated software security testing. Mobile app vetting solutions, exemplified by Kryptowire's products, can be one component among several that can work together to improve the security of the mobile app ecosystem and supply chain.

Acknowledgements

We are grateful to the following individuals for their generous contributions of expertise and time in conducting this pilot and report.

Name	Organization
Vincent Sritapan	Department of Homeland Security, Science and Technology Directorate
Michelle Brown	National Information Assurance Partnership
Mary Baish	National Information Assurance Partnership
Robert Clemons	National Security Agency
Kevin Gallicchio	National Security Agency
Joseph McDaniels	National Security Agency
Zachary Smith	National Security Agency
Chris Gogoel	Kryptowire
Amit Sharma	Leidos
Kevin Steiner	Leidos
Matthew Stern	Intelligent Waves
Daniel Faigin	The Aerospace Corporation
Sheldon Durrant	The MITRE Corporation
Terri Phillips	The MITRE Corporation
Carolyn Francisco	The MITRE Corporation

Table of Contents

Table of Figures

1 Introduction and Purpose

The federal government's increased use of mobile devices and mobile applications (apps) and the need to provide assurance of the security of those apps, have resulted in a market of mobile app vetting solution providers. Application software with cybersecurity functionality in National Security Systems (NSS) must be evaluated against the security requirements defined in the National Information Assurance Partnership (NIAP) Protection Profile (PP)[1] for Application Software. Over the last few years, the number of providers whose products include testing against security requirements specified in the NIAP application software PP has increased.

While these commercial app vetting products include tests for NIAP requirements, their utility to support mobile app software evaluations conducted by NIAP-approved Common Criteria Testing Laboratories (CCTLs) has not been explored. In 2018, the Department of Homeland Security (DHS) Science and Technology Directorate (S&T), in partnership with subject matter experts from the National Security Agency (NSA), NIAP validators, and Leidos Common Criteria Testing Laboratory, initiated a pilot using Kryptowire's software assurance tool to automate testing against PP requirements for the Hypori Virtual Mobile Infrastructure client software, a virtual smartphone solution.[2,3,4] The automated requirements testing pilot was funded by DHS S&T as part of its ongoing Mobile Security Research and Development program.

Mobile apps are created and updated more frequently than software deployed to desktops and servers. While the NIAP evaluation process can be completed in just 90 days, it may take up to six months,[5] which can be a costly and time-consuming process considering the length of the mobile app development cycle.

The goal of this pilot was to determine how much of the NIAP evaluation testing could be automated, thereby improving efficiency, increasing throughput (number of apps evaluated), and reducing cost of NIAP testing. This report discusses the findings and results of this partnership effort, feedback and lessons learned from the participants. It also proposes next steps to promote adoption of the automated approach.

[1] NIAP PPs specify an implementation-independent set of security requirements for a category of information technology products that meet specific federal customer needs. The NIAP PPs are intended for use in certifying products for use in National Security Systems to meet a defined set of security requirements; NIAP PP certified products are also used by federal organizations in non-National Security Systems.

[2] "DHS S&T Awards SBIR Contract to Mclean Small Business for Mobile Security Research and Development." DHS S&T. July 12,2016. https://www.dhs.gov/science-and-technology/news/2016/07/12/news-release-st-awards-750k-sbir-contract-mclean-company.

[3] "Intelligent Waves Awarded $43 Million Small Business Innovation Research Contract." Intelligent Waves LLC. December 18, 2019. https://intelligentwaves.com/2019/12/19/innovation-research-contract/.

[4] NIAP Product Compliant List: Hypori Client. https://www.niap-ccevs.org/Product/PCL.cfm?par303=Intelligent%20Waves%2C%20LLC.

[5] https://www.niap-ccevs.org/Ref/Evals.cfm

1.1 Background

1.1.1 National Information Assurance Partnership

NIAP is responsible for U.S. implementation of the Common Criteria, including management of the Common Criteria Evaluation and Validation Scheme (CCEVS) validation body. NIAP manages a national program for developing protection profiles, evaluation methodologies, and policies that will ensure achievable, repeatable, and testable requirements. It also participates in international standards bodies and working groups to create common security requirements and methodologies so products developed and evaluated outside of the U.S. can be used to meet U.S. government needs. NIAP, through the National Institute of Standards and Technology (NIST)-administered National Voluntary Laboratory Accreditation Program (NVLAP), also approves commercial Common Criteria Testing Laboratories to conduct these security and cryptographic evaluations.

1.1.1.1 NIAP Protection Profiles

NIAP PPs specify an implementation-independent set of security requirements for a category of information technology (IT) products that meet specific federal customer needs. NIAP security evaluations are conducted by approved independent commercial testing laboratories. For this pilot, the Intelligent Waves Virtual Mobile Infrastructure Platform Hypori client app was evaluated against the PP for Application Software, version 1.2.

1.1.1.2 NIAP Sponsor/Developer

The NIAP sponsor may be a product developer, a value-added reseller of an IT security-enabled product, or another party that wishes to have a product evaluated. The sponsor hires and works with a CCTL to conduct a security evaluation of an IT product. Intelligent Waves was the sponsor for this evaluation of the Hypori client (Android and Apple iOS) portion of Intelligent Waves' Virtual Mobile Infrastructure Platform.

1.1.1.3 Common Criteria Test Lab

The CCTL is a commercial testing laboratory accredited by NIST's NVLAP and approved by NIAP to perform security evaluations against the *Common Criteria for Information Technology Security Evaluation* using the assurance activities defined in one or more protection profiles, and, where appropriate, the procedures defined in the *Common Methodology for Information Technology Security Evaluation.* CCTL evaluators perform a variety of activities for each evaluation, including analyzing a product's security claims, providing consulting services to product vendors, performing security testing on the product, and documenting the results of the evaluation. Leidos conducted the NIAP evaluation of the Hypori client apps.

1.1.1.4 Cryptographic and Security Testing Lab

Strong cryptography is a core part of the security of many products, whether on its own or when used as part of other capabilities such as secure communication protocols. A Cryptographic and Security Testing (CST) laboratory is a commercial testing laboratory accredited under the auspices of NIST's NVLAP to perform testing of products that implement cryptographic functionality. CSTs are accredited to conduct testing of two kinds of cryptographic components:

those that implement cryptographic algorithms per NIST's *Cryptographic Algorithm Validation Program*, and cryptographic modules per NIST's *Cryptographic Module Validation Program*. The mobile platforms on which the Hypori clients were installed and tested were products consisting of NIST-evaluated cryptographic algorithms, and cryptographic hardware and software modules.

1.1.1.5 NIAP Validators

A validation team is assigned to each evaluation to act as independent technical liaisons between NIAP and the CCTL and to ensure the evaluation meets NIAP standards and satisfies the requirements of the Common Criteria Recognition Arrangement (CCRA). The validation team advises the CCTL on both technical and process issues but does not produce evaluation evidence. Upon conclusion of each evaluation, the validation team reviews the test evidence and documentation produced by the product vendor and CCTL, evaluates the adequacy of the testing, and assesses whether the evaluated product meets all requirements of the applicable PP(s).

1.1.1.6 NSA Subject Matter Experts

The NSA provides subject matter experts (SMEs) in cybersecurity and specific technology areas to support the NIAP program. These SMEs serve a variety of roles including providing input into the development of the NIAP-approved PPs in various technology areas, serving as members of various working groups and technical communities where security requirements are developed for inclusion in the PPs, evaluating specific aspects of the evidence provided by the vendors and CCTLs , and serving as advisors to adjudicate issues that arise regarding how to interpret, test, or analyze the results of evaluation evidence.

1.1.2 DHS Science and Technology Directorate

The DHS S&T Office of Mission and Capability Support (MCS) works with DHS operational Components, first responders at all levels of government, emergency management and public safety personnel, and other homeland security organizations to define priorities, gaps, and requirements to find or develop technology solutions. S&T MCS created the Mobile Security Research and Development (R&D) Program, with the goal: "Accelerating the adoption of secure mobile technologies by government and industry to enable the homeland security mission." The Mobile Security R&D Program funded Kryptowire LLC, a mobile app vetting solution provider, to automate mobile app vetting based on NIAP standards. DHS S&T provided the funding and technical support for Kryptowire to evaluate the Hypori client apps against the Application Software PP.

1.1.3 Kryptowire LLC

Kryptowire tests mobile and Internet of Things firmware and apps against the internationally recognized security standards used for classified and NSS. Kryptowire continuously assesses the security of all enterprise mobile apps and devices against the software assurance standards published by NIST, NIAP, and the Open Web Application Security Project Top Mobile Security Risks. Kryptowire used its software to test the Hypori clients against the NIAP standards and provided the output to the CCTL and NIAP validators.

1.1.4 Intelligent Waves

Intelligent Waves' Hypori is a proprietary virtual smartphone technology that virtualizes the entire mobile experience—no data or applications reside on the device. DHS S&T awarded Hypori a Small Business Innovation Research Phase III contract. With the contract, Intelligent Waves will provide a pre-production implementation of the Hypori Virtual Mobile Infrastructure (VMI) capability for DHS S&T-sponsored government customers to conduct Hypori-as-a-Service end-user evaluations. The Hypori client app portion of Intelligent Waves' VMI was tested for the automation pilot.

2 The NIAP Evaluation Process

2.1 NIAP Evaluation Process Goals

The NIAP evaluation and accreditation process is designed to provide some degree of assurance that commercial off-the-shelf (COTS) products that are procured for use within NSS meet specific security standards and requirements. The evaluation of a product, known as the Target of Evaluation (TOE), against the NIAP Protection Profiles produces the following three kinds of evidence:

1. Those that describe how the product implements the security requirements.
2. Those that evaluate the documentation that accompanies a product to ensure it describes how administrators should configure the product to meet the security requirements.
3. Those that test the functionality of the product to ensure it meets the security requirements.

NIAP implemented the evaluation process with several objectives in mind. These include, but are not limited to the following:

- **Transparency:** End-users and other stakeholders interested in the evaluation and evaluated products should be able to gain insight into what security claims were made, whether the product met the claims, how the product met the claims, and whether there were any issues related to the product of which they should be aware.
- **Repeatability:** The requirements, evaluation activities, testing methodologies, and evaluation evidence should be clear, consistent, coherent, accurate, and technically sound enough to enable one to repeat the evaluation activities and arrive at the same conclusion.
- **Timeliness:** The end-to-end evaluation of a given product should not take overly long; the process should be fast enough to enable the government to get timely access to the products they need to fulfill their mission objectives.

2.2 NIAP Evaluation Artifacts

The evaluation process requires the independent testing laboratories and vendors to conduct a variety of activities and to produce a variety of documents, including but not limited to the following:

- **Security Target (ST):** This document contains the set of security claims against which the product is to be evaluated. It also contains a description of how the product implements and meets the security requirements from the protection profiles (documented as part of the TOE Summary Specification [TSS]). The ST and TSS are publicly available documents.

- **Entropy Assessment Report (EAR):** This report describes how the product implements the underpinnings of core cryptographic functionality, including the source(s) of entropy used to generate random numbers, cryptographic keys, etc. It provides an assessment and set of assertions to ensure that enough entropy reaches the system components responsible for cryptographic functionality. The EAR may also contain a key hierarchy document that describes how and where cryptographic keys are generated and used and their relationship to each other. It is a proprietary document not available to the general public.
- **Detailed Test Report (DTR):** The DTR documents detail the testing environment, test steps, test cases, pre-conditions, post-conditions, expected results, actual results, and evidence gathered for each test to which the TOE is subjected. It is a proprietary document not available to the general public.
- **Assurance Activity Report (AAR):** The AAR includes a summary of how the TOE meets each requirement and includes evidence pulled from the ST/TSS, DTR, and related administrator and user guidance documents. The AAR is a publicly available document.
- **Validation Report (VR):** The VR documents the activities that took part during the product evaluation and includes NIAP's assessment and conclusions that the TOE has successfully completed evaluation. It also includes any issues identified during the evaluation effort that merit special attention by the product's end-users. The VR is a public document.

2.3 NIAP Evaluation Process Challenges

Although NIAP has made strides in recent years to vastly improve the viability, usefulness, speed, and the number of products approved for use to support the government's national security missions, challenges and areas for improvement remain. These include timeliness and cost effectiveness, completeness and accuracy, and consistency.

2.3.1 Timeliness and Cost Effectiveness

NIAP continues to address timeliness and cost challenges for evaluation of COTS products.

COTS products are developed and sold in an environment in which speed to market is often imperative to a vendor's market success. If the time and cost needed to complete NIAP certifications can be reduced, users of NIAP-certified products can take advantage of emerging trends and current technology, thereby avoiding obsolescence. Reductions in certification time and cost also address vendor concerns about maintaining older certified products specifically to accommodate a government user base. The use of older products presents security concerns in that older products often have unpatched vulnerabilities or do not take advantage of security architecture improvements that are available in newer products. NIAP introduced its Assurance Continuity program to improve the certification process to address some of these concerns and continues to search for additional ways to reduce time and cost.

2.3.2 Completeness and Accuracy

Evaluations of the security-relevant features implemented in IT products can be deemed credible only if they completely and accurately capture evidence that enables stakeholders to determine whether the features meet the product developer's claims. A combination of factors contribute to achieving accuracy in the evaluations. These include the specification of clear, objective

requirements; a clear understanding of acceptable and unacceptable testing methods and approaches; access to tools and technologies that enable evaluators and certifying bodies to understand and assess how the product works and gather the relevant supporting evidence; and documenting the evidence, analysis, and verdicts appropriately.

2.3.3 Consistency

NIAP administers the U.S.'s implementation of the CCRA, an international agreement among several nations so that products evaluated in one country are mutually recognized by the other member nations without the need for retesting and recertification in each participating nation. Among the CCRA's objectives is that product evaluations be conducted to consistent standards so participating nations can have confidence in the security of the evaluated products. Common Criteria members accomplish this consistency by specifying a framework for the development and testing of objective security requirements, ensuring that testing labs are accredited to common standards, and auditing the framework periodically to ensure its quality and adherence.

2.4 Test Automation

A typical NIAP evaluation consists of testing the TOE against the Security Functional Requirements (SFRs) defined in one or more PPs and for which the product vendor claims support. Each PP may contain dozens or hundreds of these SFRs, which in turn may require one or more test activities that prove whether the TOE fulfills the requirements. Test activities produce a variety of evidence and artifacts, including, but not limited to, system log files, packet captures, images, video, raw system data, and text. CCTL evaluators gather this evidence using a variety of manual and automated means, often using elaborate testbeds and sophisticated hardware and software tools. This test evidence is compared and evaluated against a large body of documentation such as the Security Target and TOE Summary Specification, the vendor's product documentation and administration guide(s), analysis of the product's entropy information, and others.

An evaluation involves the production and analysis of a significant amount of documentation and evidence, and considerable time and effort.

All told, even a relatively simple evaluation involves the production and analysis of a significant amount of documentation and evidence, and considerable time and effort. Not surprisingly, there is a natural push to improve the efficiency of the overall process by turning to automation. Many of the CCTLs already have large, complex testbeds that are used to evaluate various SFRs claimed by a TOE. However, as the evaluations themselves become more complex, there is a renewed push for even greater automation of all aspects of the process by evaluation stakeholders. This push includes the automated mechanisms to specify Protection Profiles and the claims made against them, improved and more comprehensive automation of requirements testing, and automation of the documentation generation and analysis. Such automation can result in more accurate, comprehensive, and repeatable testing, while reducing overall evaluation time, improving consistency, and even revealing previously unknown issues. This report documents the findings of one such effort funded and managed by DHS S&T in partnership with NIAP to determine whether and how much of a mobile app could be evaluated using automation.

3 Pilot Overview

The DHS S&T automated app software evaluation pilot sought to determine the degree to which the Kryptowire test tool could be used to automate testing against the NIAP *Protection Profile for Application Software Version 1.2*, the applicability of its results as evaluation evidence, and if the Kryptowire tests met the intent of the PP.

Prior to the start of the pilot, Intelligent Waves requested a NIAP evaluation of its Hypori Virtual Mobile Infrastructure Version 4.1 client apps for Android and Apple iOS devices. Leidos conducted the evaluation using its defined test plans, tools, and procedures. The NIAP-assigned validation team monitored the activities of the Leidos evaluation team, examined evaluation evidence, and reviewed the evaluation results. Based on the results, the validation team determined the product satisfied the security requirements and the Hypori clients for iOS and Android were placed on the Product Compliant List in August 2018.[6]

3.1 Pilot Approach

The automation pilot with Kryptowire was conducted separately from the Leidos evaluation of the Hypori client software. It included the involvement of Leidos staff, NIAP validators, and NSA SMEs to review the findings and evidence and provide feedback to Kryptowire regarding applicability of its reports as evaluation evidence. The approach consisted of three rounds of reporting, feedback, and updates to Kryptowire's reporting and production of testing evidence as follows:

Iteration 1:

- Kryptowire scanned the Hypori client apps and provided the results to Leidos.
- Leidos reviewed the results and provided feedback on whether the results satisfied the security functional requirements.

Iteration 2:

- Kryptowire updated the reporting function and language, reran the tests in response to Leidos' feedback, and provided updated reports and evidence to Leidos.
- NIAP validators reviewed the Iteration 2 reports and provided feedback on: a) what additional items would be needed for the results to be accepted by a validation team (NIAP validator review), and b) if the tool results and findings met the intent of the SFRs (NSA SME review).

Iteration 3:

- Kryptowire made additional improvements based on Iteration 2 comments from Leidos and the NIAP validators and NSA SMEs reviews.
- NIAP validators and NSA SMEs provided their final comments and feedback to Kryptowire regarding sufficiency of evidence for the SFRs as written and whether the results satisfied the intent of the requirements (i.e., if the requirements were rewritten).

[6] https://www.niap-ccevs.org/Product/PCL.cfm?par303=Intelligent%20Waves%2C%20LLC

3.2 Overview of the Protection Profile for Application Software

The *Protection Profile for Application Software* (App PP) defines the security requirements that need to be met by application software that runs on mobile devices, desktops, and servers. Like many other PPs, the App PP can be used either on its own or in conjunction with other sets of requirements as defined in extended packages for more specialized types of applications. The App PP defines a software application broadly to cover a variety of scenarios, including cases where the app is installed directly on an operating system, an execution environment running on an operating system, or a combination of the two. It also includes any kernel-resident code that is installed as part of the application package and is required for the app to function and to meet the SFRs.

The App PP itself covers a variety of threats to software security, including network-based attacks, eavesdropping, localized attacks to the application's platform, and attacks that require physical access to the platform. It specifies requirements to mitigate these kinds of threats by defining security objectives for the TOE such as software integrity, software quality, management of software security functions, provisions for user and software secure data storage, and protected communications.

The App PP recommends that some security-critical functions be delegated to the application's execution platform, e.g., Microsoft, Apple, Android. Test assurance activities vary depending on the platform.

Like many other PPs, the App PP recommends that certain security-critical functions be implemented by the platform on which the application executes. This recommendation is offered because relying on the platform for core security functionality reduces the likelihood that app developers will produce their own flawed implementations, the platform's functionality can be evaluated exhaustively once and then approved and reused for different classes of applications, any flaws discovered can be fixed in the platform, and applications can take advantage of the fixes with little or no effort. However, the App PP is written to be flexible so software developers can provide their own implementations of security requirements, along with assurance activities that the developers' own implementations must meet to ensure compliance. For this reason, many of the requirements allow the application developer to specify whether a given requirement is met exclusively by the app, exclusively by the platform, or by a combination of both. Besides the mandatory requirements, the App PP also provides a set of requirements, known as selection-based requirements, that must be met depending on whether related requirements are claimed by the developer; optional requirements, which the TOE developers may claim, depending on whether the product implements certain features; and objective requirements, which capture functionality that is not currently mandated, but will be in the future, and whose adoption is therefore encouraged.

The App PP is relatively unique among PPs in that for many requirements it defines test activities that vary depending on the platform on which the app executes. For example, it defines a set of tests that an evaluator should execute to provide evidence that an app developed for Microsoft platforms is compliant and a corresponding set of tests for other platforms such as Android, Linux, or Apple platforms. The App PP version against which the Hypori apps were tested consists of the following categories of SFRs:

- **Cryptographic Support:** Several SFRs that cover such cryptography-related aspects as random bit generation; key generation and management; cryptographic operations such as encryption and hashing; and use of digital certificates.
- **User Data Protection:** Includes protection of user-level data and policies covering access to sensitive hardware features such as network connectivity, camera and microphone, location services, and other communication radios; also includes encryption of data at rest using file, folder, or full-disk encryption.
- **Security Management:** Covers secure configuration management as well as the specification of management functions.
- **Privacy:** Covers the transmission of sensitive user data such as Personally Identifiable Information (PII).
- **Protection of the TOE:** Covers mechanisms designed to protect the TOE from malicious activity such as the following items: which platform Application Programming Interfaces (APIs) are approved, implementation and use of anti-exploitation capabilities, software integrity and secure software updates, and use of third-party libraries.
- **Protection of Data in Transit:** Includes requirements for the establishment and use of secure data transmission protocols such as Transport Layer Security, Hypertext Transfer Protocol Secure, and Secure Shell.

4 Pilot Results and Findings

The results of the pilot were both illuminating and surprising in several ways. While its primary objective was to investigate whether and how much of the testing against the App PP could be automated in its current form, the test pilot also provided valuable insights into the efficacy of the App PP's assurance activities and test procedures as well as the entire evaluation process itself.

As shown in Figure 2, during the test pilot Kryptowire analyzed both the iOS and Android versions of the Hypori application by conducting iterative runs of the Hypori application against their scanning tools. Kryptowire provided multiple reports and related test artifacts to the Leidos evaluation team, who had conducted a traditional evaluation of the product using their own testbeds, automated and manual testing procedures, and their own customary evaluation methodologies. The Leidos evaluators compared the results of their testing approach with the Kryptowire generated results and provided comments and their assessment of how well the automated Kryptowire testing and test evidence met the requirements. The final Kryptowire report and test evidence was provided to NIAP validators and NSA SMEs for their input and analysis.

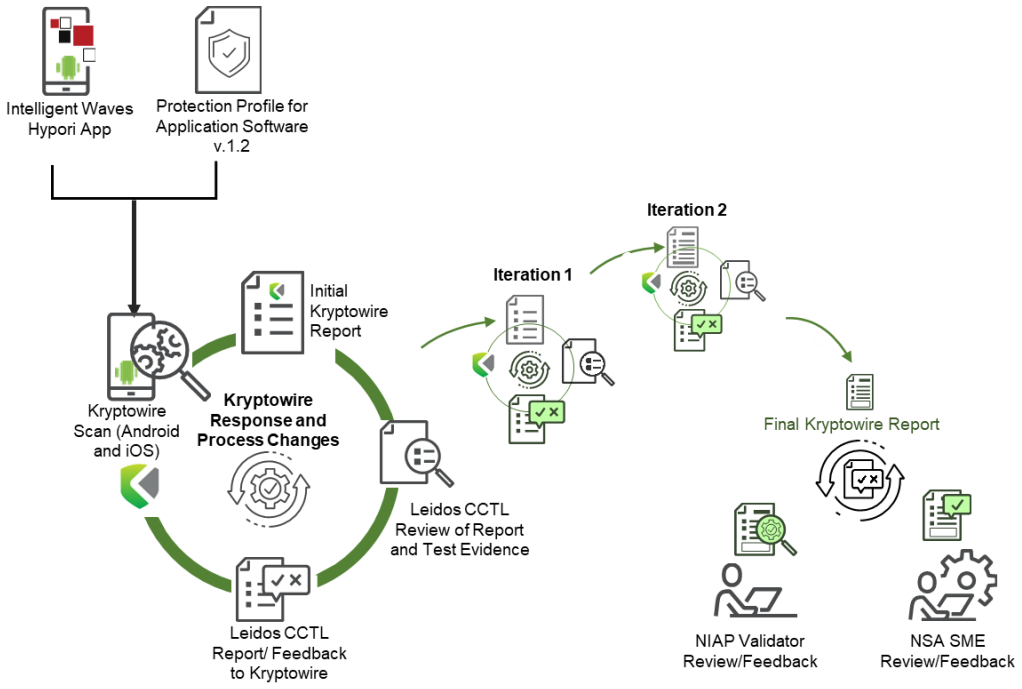

Figure 2. NIAP Automation Pilot Process

There are a few noteworthy items regarding Kryptowire testing versus a traditional testing approach. First, Kryptowire's initial set of tests were not conducted against the requirements as documented in the Hypori Security Target because it did not have access to these requirements. Typically, before an evaluation is conducted, the vendor or the testing lab creates a Security Target specific to the TOE by including mandatory, selection-based, optional, and objective SFRs that the TOE implements. Kryptowire targeted all 25 mandatory SFRs in the PP and assumed no knowledge of the selections made in the Security Target. In some cases where relevant data already existed, Kryptowire added selection-based requirement data (one for

Android, three for iOS) in the last iteration of testing. The Kryptowire testing also did not account for the application of any Technical Decisions, which are modifications made to a PP's requirements to correct errors, clarify requirements, or address other issues with the PP.

The following sections summarize some findings from the test pilot.

4.1 Example Findings

Leidos CCTL evaluators performed several analyses of Kryptowire's results and associated test evidence for both Android and iOS. Iteration 1 (the initial Kryptowire scan) and Iteration 2 (which included responses and updates based on feedback from Leidos) were deemed to have many inconclusive results, which Kryptowire addressed through several product and process updates. The report generated for the final iteration (Iteration 3) was provided to NIAP validators and NSA SMEs for additional input and analysis, particularly in response to the results that Kryptowire flagged for additional review (31 percent for Android and 41 percent for iOS). Kryptowire assigned this status for those results requiring analyst review to make the final determination regarding whether the result should receive a pass (69 percent for Android and 55 percent for iOS) versus fail (4 percent for iOS) score after reviewing the raw data gathered by the Kryptowire product.

Kryptowire's Iteration 3 results are summarized in Figure 3:

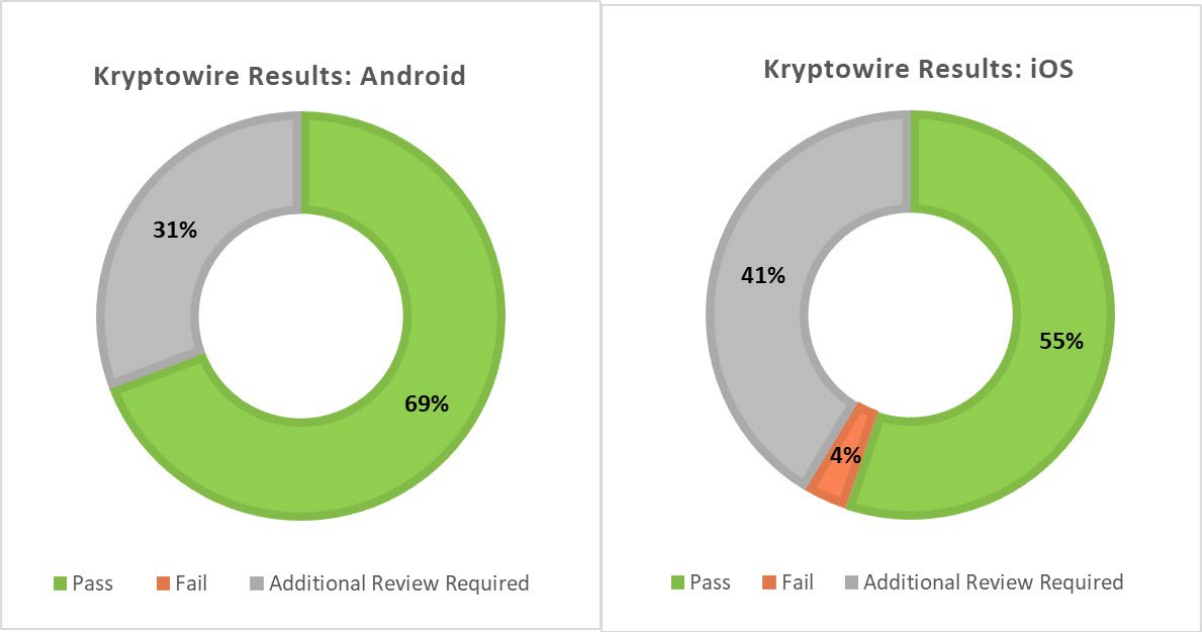

Figure 3. Kryptowire's Results for Android and iOS.

NIAP's additional analysis concluded that Kryptowire's testing was indeed an accurate reflection of Hypori's ability to meet the SFRs. NIAP's analysis revealed the following:

- Kryptowire's initial testing was not conducted against the security claims specified in the Hypori Security Target; rather the tests were taken directly from the App PP. This approach resulted in a high number of false-positives because the App PP includes all possible requirements, including those that are optional, selection-based, and objective. Therefore, many of the results that Leidos determined were inconclusive would not have applied to the

Hypori application. As Kryptowire's testing incorporated more of the activities and data that are characteristic of the way traditional test labs conduct their evaluations, Kryptowire's automated testing became more consistent with Leidos' conclusions.

- The NSA SMEs concluded that many of the discrepancies stemmed from the fact that while the Kryptowire tools, as implemented, did not meet the letter of the testing assurance activities documented in the App PP, they nevertheless met the intent of the requirements. This result is due in part because the assurance activities specify a method or approach to prove that the TOE meets the SFRs, but these are not necessarily the only ways to prove the TOE's compliance. For example, one could test that the TOE uses the appropriate ciphers in secure protocols by iteratively establishing a secure connection using each implemented cipher suite and verifying via analysis of network traffic that only the approved ones result in successful connections. Alternatively, one could conduct static and/or dynamic analysis on the TOE's application binary to determine which ciphers are implemented and executed when the TOE establishes secure connections. Both approaches produce strong evidence that the TOE is compliant with the requirements, however, the second approach requires different tools, skill sets, and analysis than the first, thereby contributing to the perception of inconclusive or inconsistent results.
- Like all evaluations, the Kryptowire testing resulted in the creation of a large amount of evidence. While Kryptowire's reports provided a concise overview of whether the TOE was compliant by pointing to specific, often highly detailed test results, the initial reports sometimes lacked the contextual richness that is typical of the reports produced by testing labs. Without this context, the personnel reviewing the test results and evidence often had difficulty determining whether the results were accurate and satisfied the requirements. Kryptowire was able to successfully modify their reporting and their presentation of test evidence in ways that improved the reviewers' ability to determine that the TOE met the intent of the requirements.

Static and analysis testing can reveal "dead code" in an app, which may not otherwise be identified during NIAP evaluation testing

- The Kryptowire tool's static and dynamic analysis capability could detect and bring to the surface issues that are not necessarily apparent by conducting testing using the methods prescribed in the PPs or using the methodologies typically used by the testing labs. For example, static and dynamic analysis can reveal the existence of "dead code" that is present in a system or vulnerabilities in the libraries that are used in an app.
- Kryptowire's tools and technology were able to accurately capture the TOE's security characteristics and functionality without requiring access to the TOE's source code and without support from the mobile device platform vendor (e.g., Samsung or Apple for Android and iOS, respectively). NIAP evaluations do not require access to source code or platform vendor support for app testing, so the pilot did not explicitly seek to demonstrate this capability. Nevertheless, one outcome of the effort was that Kryptowire demonstrated that such access and support are not needed to support new and emerging use-cases and operating models that take advantage of NIAP PP requirements and methodologies. This means, for example, that government agencies can assess prospective apps for PP compliance (including those downloaded directly from official app stores) without necessarily needing to enlist the help of external parties (e.g., the original app developer).

To summarize, Figure 4 shows that NSA SMEs found the Kryptowire results detailed and accurate and assessed whether the findings met the intent of the requirement. After reviewing Kryptowire's test results and supporting evidence, the NSA SMEs determined that almost 90 percent of the results met the intent of the requirements for iOS, and more than 60 percent for Android.

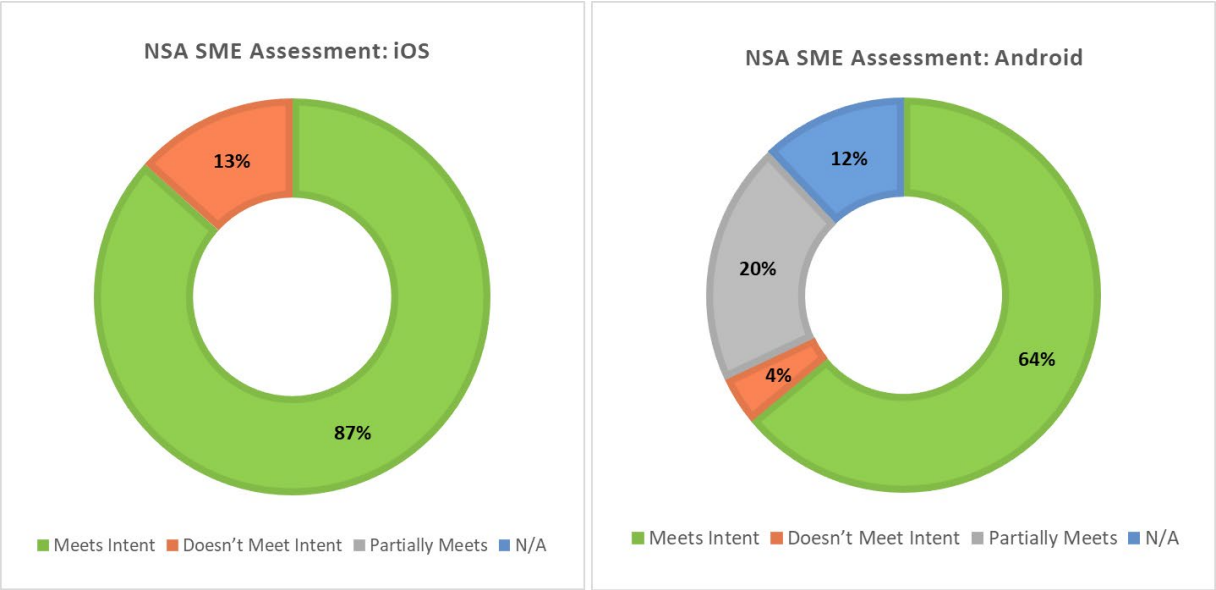

Figure 4. NSA SME Assessment of Whether Kryptowire's Evidence Meets Intent of Requirements.

5 Recommendations and Conclusion

DHS S&T's automation test pilot evaluation of the Hypori application using Kryptowire's app security software was successful. It not only proved that it is possible to use Kryptowire and similar automated tools as a major part of a NIAP evaluation, but also that the results can be relied upon to be accurate and trustworthy. The pilot also proved that it is possible to use the tools to quickly and consistently gather test evidence that analysts and NIAP validators can use to determine whether a product meets the security requirements specified in the App PP. While manual testing in traditional evaluations may take days or weeks, automated app evaluation can be completed in hours or even minutes. Figure 5 summarizes performance statistics gleaned from scanning the top 100 free apps available on the market from the Google Play store and the Apple App Store over a one-year period. Even when differences in size and complexity of the apps are considered, one can see that it takes relatively little time to complete an automated analysis. Such efficiency gains benefit the overall evaluation process in a variety of ways. It allows app developers to quickly identify and address potential issues in their products prior to formally entering the NIAP evaluation process. It allows analysts and evaluators to gather raw data from an app for further analysis and documentation as part of the formal evaluation, and, if needed, for further testing and remediation. This automated testing also integrates well with the rapid, iterative, agile development and release cycles characteristic of modern software development, allowing successive versions of an app to be evaluated and approved quickly. These efficiencies also drive down evaluation costs because it reduces the amount of time and the number of personnel needed to complete a certification for all stakeholders.

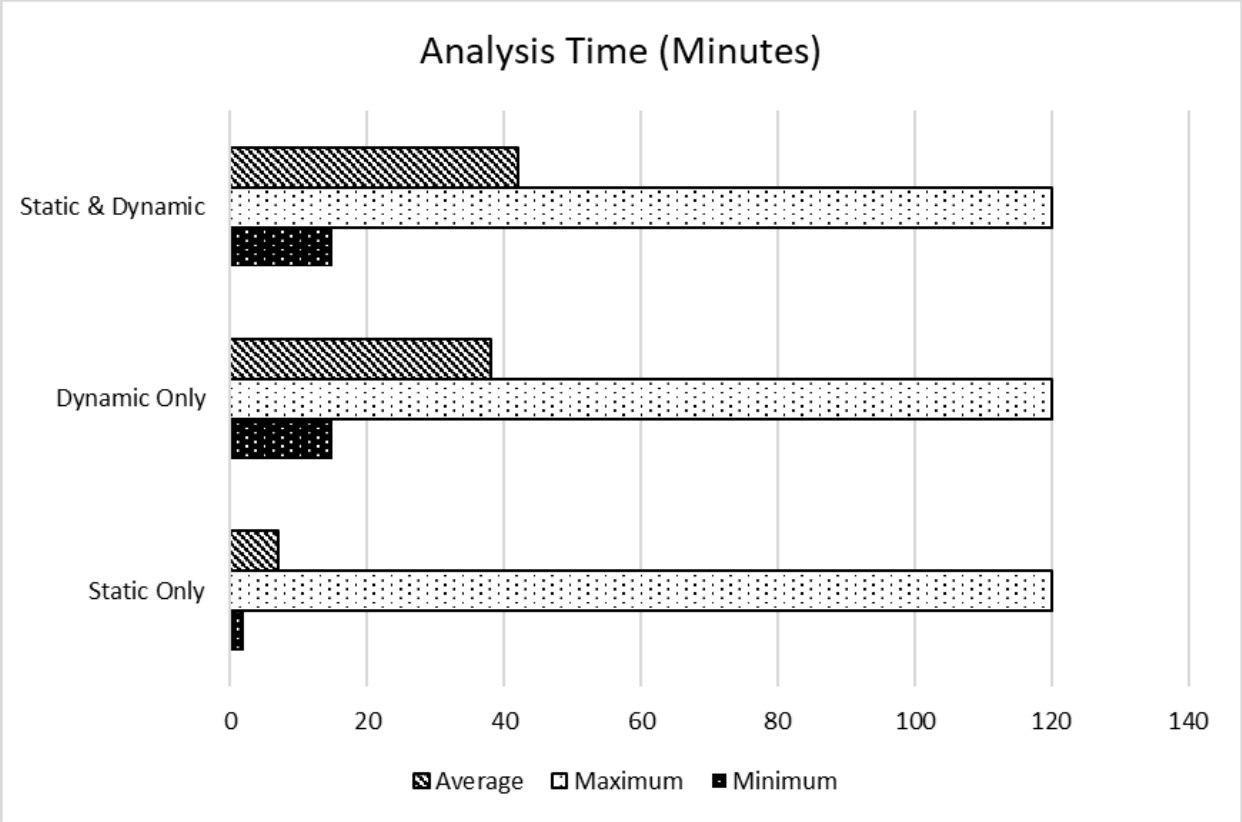

Figure 5. Analysis Time for Automated App Analysis.

While the pilot was successful, it did expose some areas for improvement in the evaluation process. For example, in a traditional evaluation, a lab's findings are typically documented in such a way as to correlate evidence from several sources, including from a variety of platform tools, data from testbed networks, vendor-provided configuration and administration guides, and other information sources. This means that while the Kryptowire tool provided extremely valuable insights and highly detailed and accurate results, it provided only a part of the evidentiary information that comprises a traditional NIAP evaluation and would therefore need to be combined with other information and documentation to meet the complete set of evaluation activities.

Another area for improvement involves the incorporation of Technical Decisions (TDs). A TD is a modification to a PP that arises from an identified problem with the requirements in a PP or a clarification of how the requirements should be interpreted. The TDs may include one or more changes and can affect the way one or more requirements are specified, interpreted, tested, and/or documented. TDs are especially challenging in that they can be issued at any point prior to the end of an on-going evaluation and all relevant TDs must be addressed by the product's developers to be certified compliant. Automation could potentially address some of these challenges regarding TDs, for example by ensuring that all relevant TDs are applied to a given product. If a late-breaking TD imposes significant requirements or testing changes, however, these tests may need to be done manually because automated tool vendors may not have enough time to analyze, design, develop, test, and deploy the changes to their respective testing environments prior to the end of an evaluation.

According to Kryptowire and NIAP, if changes were made to how requirements, assurance activities, and acceptable evidence are specified in PPs, it may be possible to automate as much as 90 percent of testing against the App PP.

The pilot also brought to light areas for improvement in the way NIAP PPs are developed and used and the assumptions for the kinds of evidence that are acceptable. It also reinforced that some of the test methods and tools currently included in the App PP are not the only way or the best way to test that a mobile app meets the requirements in the App PP. According to Kryptowire and the NIAP, if changes were made to how the requirements, assurance activities, and acceptable evidence are specified in the PPs, it may be possible to automate as much as 90 percent of the testing against the App PP.

There are already efforts underway within NIAP and the wider community to improve various aspects of the App PP, including allowing a greater variety of test evidence to be provided as well as to make automated testing easier, faster, more consistent, more accurate, and more conclusive. This work includes the specification of machine-readable security requirements and security claims, the development of tools to better capture and document test evidence, and data exchange protocols to speed the transfer and analysis of test evidence. As these developments continue, additional benefits of automation could be realized such as reducing evaluation costs by allowing a mobile app to be subjected to periodic retesting and recertification, pursuant to the NIAP Assurance Continuity process, as developers release new versions. App developers also benefit from the ability to determine ahead of time whether their apps are compliant to the NIAP standards before submitting them for official certification, which reduces both costs and risk, especially for smaller businesses or those with limited financial and personnel resources. As a result, a much larger number and variety of apps can be certified for use in a shorter period and the entire certification process can be more responsive to evolving security threats.

Automation provides a fair amount of risk reduction for federal agencies as well. The authorizing officials within federal agencies benefit from increased automation because it helps to drive a more seamless, cost-effective process toward full NIAP mobile app vetting certifications. Automated testing can also be used to provide a middle ground to demonstrate compliance with the NIAP requirements pending a full certification, thereby enabling authorizing officials to provisionally make an immediate risk determination to use an app in the short term until a full certification is completed.

Automated evaluations against requirements specified in NIAP PPs also provide authorizing officials additional flexibility. Officials may, for example, quickly determine which commercially available apps are compliant with the App PP even if those apps will not be submitted for an official NIAP certification, as may be the case for apps that are acquired from public app stores. The overall effort demonstrates these possibilities and aims to make the ecosystem stronger, thereby enabling authorizing officials to use mobile apps securely for their respective agency's missions. It therefore behooves stakeholders from both government and industry to continue to develop and prove the capability of automated testing tools and methodologies to address the increasing scale and complexity of certifying secure mobile apps for use within the federal government.

Appendix A Analysis Reports

NIAP Product Compliant List Reports

Prior to the automation pilot, Intelligent Waves requested a NIAP evaluation of its Hypori VMI Version 4.1 client apps for Android and Apple iOS devices. Leidos conducted the evaluation; the NIAP-assigned validation team monitored the activities of the Leidos evaluation team, examined evaluation evidence, and reviewed the evaluation results. Based on the results, the validation team found that the product satisfied the security requirements and the Hypori clients for iOS and Android were placed on the NIAP Product Compliant List (PCL) in August 2018. Reports can be found at:

- Compliant Product - Hypori Client (iOS) v4.1
- Compliant Product - Hypori Client (Android) v4.1

Kryptowire NIAP Compliance Reports

Kryptowire generated reports and related test artifacts for review by the Leidos evaluation team, NIAP validators, and NSA SMEs. The final iteration of the NIAP analysis reports can be found at: https://www.kryptowire.com/niap-dhs-study/

- Kryptowire Android NIAP Analysis, Hypori client app version 4.1.5
- Kryptowire iOS NIAP Analysis, Hypori client app version 4.1.3

List of Acronyms

Acronym	Definition
AAR	Assurance Activity Report
CCEVS	Common Criteria Evaluation and Validation Scheme
CCRA	Common Criteria Recognition Arrangement
CCTL	Common Criteria Testing Laboratory
COTS	Commercial-off-the-Shelf
CST	Cryptographic and Security Testing
DHS	Department of Homeland Security
DTR	Detailed Test Report
EAR	Entropy Analysis Report
IT	Information Technology
MCS	Mission and Capability Support
NIAP	National Information Assurance Partnership
NIST	National Institute of Standards and Technology
NSA	National Security Agency
NSS	National Security Systems
NVLAP	National Voluntary Laboratory Accreditation Program
PCL	Product Compliant List
PP	Protection Profile
R&D	Research and Development
S&T	Science and Technology Directorate
SFR	Security Functional Requirements
SME	Subject Matter Experts
ST	Security Target
TD	Technical Decision
TOE	Target of Evaluation
TSS	TOE Summary Specification
VMI	Virtual Mobile Infrastructure
VR	Validation Report