Intel Science and Technology Center for Secure Computing
University of California at Berkeley: Secure Computing Research for User Benefit

Overview

Computer technology has enriched billions of people's lives by vastly expanding the volume, quality and accessibility of information, simplifying communication and collaboration, enabling the production, distribution and use of entertainment, improving the quality of government services, controlling physical infrastructure more efficiently, expanding commerce, and even fostering a new domain of "digital commerce." It is also improving medical and educational services and dramatically expanding our scientific knowledge, as well as our ability to apply this knowledge to further improve people's lives. Misuse of computer technology by adversaries may severely limit these benefits by threatening the safe use of technology in increasingly critical aspects of our lives. The Intel Science and Technology for Secure Computing community is devoted to making computer technology safe and effective for its users. This goal is emphasized by our center's slogan "Secure Computing Research for User Benefit."

Security is a complex topic and having a significant and positive impact for real users requires our effort be thoughtful and balanced approach that addresses many issues. This complexity is caused by a number of factors.

First, security is not indivisible. For example, reading the New York Times on a computer does not generally involve a critical activity in and of itself; however, if reading the New York Times on a computer places other critical activities (banking, healthcare, commerce) at risk by providing an opportunity for an adversary to download *malware* during browsing, the risk calculus changes.

Second, determining the value of security and, concomitantly, the cost benefit of a security measure is challenging. Ultimately, the users themselves must judge the level of risk they are willing to tolerate along with the cost they are willing to bear in reducing risk. This is reflected in an old security bromide: "the enemy of good security is the requirement that security must be perfect." In the physical world, the threat of being struck by lightning is real but investing in extraordinary measures of protection is generally not considered worthwhile: installing lightning arrestors and avoiding dangerous locations in a storm are probably adequate to insure a very high degree of safety at quite modest cost. Of course, making such an assessment requires that a person (user) be able to adequately gauge the residual risk and the cost of a proposed solution.

Unlike the physical domain where consequences are often immediately and clearly visible and the cost of recovery is quite clear, successful cyber-attacks are often hidden and awareness of consequences is delayed. Thus, security technology must help users be aware of actual risk[1], and be knowledgeable in the impact of protective measures, as well as their costs. With this ability and the availability of cost effective protective measures and techniques, users can enjoy the tremendous benefits of the cyber world with the same confidence (and hopefully better safety!) as the physical world. Today, users are faced with difficult choices and are armed with confusing solutions that only partially address these needs under the best of circumstances.

While many of these aspects of computer security have developed over and existed for many years, two developments have greatly amplified the importance of computer security.

The first, exemplified by the widespread use of cloud computing by consumers and corporations as well as the proliferation of mobile devices, is the central, pervasive position computers have on our lives: we communicate, purchase, are entertained by, work on and learn on digital devices which are interconnected. Users want all their devices to share their information (safely). Corporations want all their employees to become more productive and innovative by leveraging seamless access to all appropriate corporate data (while preventing competitors from having access to that corporate data). Service providers want their customers to have seamless access to their services, including services, which, for maximum benefit, access personal information on behalf of the user. Increasingly, all sensitive information is shared electronically, frequently in cloud datacenters. Arguably, for the first time in the computer industry, lack of security is the primary barrier to product adoption,

Second, the Internet has seen the emergence of increasingly powerful, widely deployed, highly motivated, well financed and highly trained adversaries. These adversaries range from “local entrepreneurs,” who sell compromised machines to bot herders for spam, to sophisticated criminal organizations that steal authentication information to enable credit

[1] Unless protective measures which unarguably carry out protective measures aligned with user goals can be instituted without specific user action.

card fraud, to nation states that use security flaws to steal intellectual property, threaten national security interests and even damage physical infrastructure.

Secure Computing for User Benefit offers a magnificent research challenge: the potential impact, on real people, is enormous, the intellectual challenge is large, the scope for contribution is vast, and the emerging technical environment offers many new opportunities.

Problems and Opportunities

> ***Symantec's 2012 Norton Cybercrime Report and McAfee's second-quarter Threats Report both say that cybercrime is enjoying exponential growth at rates neither company has seen before.***

One may properly wonder if computer security is all that important, since it has often been done badly in the past. Our answer is a resounding yes. One reason is that digital devices touch every aspect of our lives and mediate much or our interaction with the world. Powerful new services, like Google Search, electronic mail, instant messaging, online banking, online entertainment (such as iTunes), online health information and government services have emerged and presented powerful reasons to enable their safe continued use.

However, we have ample evidence of security failures that threaten digital devices.

1. Malware: Among the most evident is the prevalence of malware. Malware is software (and increasingly hardware) introduced on user systems usually without user knowledge. Malware has been used to *abuse consumer devices to send spam, steal intellectual property, steal private user data, facilitate identity theft, contaminate critical data and even remotely destroy physical property*
2. Eavesdropping and data misuse: Corporations and people lose confidential data leading to embarrassment or, in cases, big financial losses.
3. Failure of critical systems: As Stuxnet demonstrated, digital devices are more and more critical to physical infrastructure and computer security failures can result in physical damage.
4. Invasion of privacy: Health information is increasingly digitized and shared. Personal preferences are collected on service sites. Personal communications, including those employing social networks such as Facebook, email and IP telephony carry many of our most private thoughts and even political opinions. Unintentional disclosure of these communications can be awkward, costly and even tragic in many cases.

Users have real questions about computer security:

1. Why is security so mysterious: I never know why I am asked to do things or why they make me more secure. I update my computer but I hear everything is still vulnerable. I don't know what's safe to do on my computer and what isn't.
2. Is there anything better than passwords: I hate passwords. If I have a few I can remember, it's not safe because if one site loses them other services I use are at risk. Malware can steal passwords and some providers have lost passwords on missing laptops and even accidentally publishing passwords on their websites. Remembering a lot of passwords is crazy. God help me if I forget my password. Besides, how does anyone know it's really me and my password from the start? My company uses PKI based authentication, which is fine when it works but Certificate authorities have been compromised and key renewal (whatever that is) is pretty unpleasant. Many web sites use CAPTCHAs, which are also painful. Somehow security gets in the way and slows things down but even with the bother, I don't feel safe.
3. What happens to my data: I use cloud services a lot and find them useful for travel, purchasing and entertainment but how do I know what happens to all my data and who can use it? It would be nice to have my medical records available digitally to my doctors but I have no idea if that is safe.
4. How can I know who or what to trust: Banks make mistakes. When something goes wrong, even if I find out, I can't tell whether something I did caused it. If I wasn't the cause, I almost never know who is or what I can do to stop it in the future. Big companies like Microsoft know what a lot of the problems are but they will never tell me the truth if it hurts their business.
5. Where is the genius bar: I have a lot of friends who think they know about security but they spend all their time on it and can't explain how it works or what I should do. I have no idea when my computer asks me if it can do something whether I should say yes or no. I want to understand enough to know what to do and why. I can do that well enough when making decisions about the safety of my house. Why can't I do that with my computer or phone?
6. How do I trust new apps: I'd like to use some of the cool apps available for my phone but how do I know they are safe? How do I know they won't ruin my phone or steal information on my phone?
7. Can I trust the computers in my car, my house: I've heard stories about Stuxnet and attacks on cars that are digitally controlled or have a lot of computers in them. If my car fails or I can't get electricity because of a computer problem that would be huge. I'd love to be able to have helpful robotic services but if the robot isn't safe because of security problems, I'd be scared to use it.
8. Can I let others use my computer if it's not busy:

While the problems are real and important, the opportunities are equally dramatic:

1. Online banking.
2. Operation of critical infrastructure.
3. Education.

4. Health-care.
5. Beneficial control of machines and physical infrastructure.
6. Communications.
7. Government services.
8. The knowledge economy.

Digital electronics has, almost uniquely, demonstrated the ability to exponentially improve (Moore's Law) and thus exponentially improve our lives. With progress in computer security, we can safely enjoy the benefits this continued progress confers as activities become safer, cheaper, more convenient, more accessible, better understood, and more reliable.

So how do we make this possible? Where are the critical opportunities based on scientific progress and collaborative research? What is our strategy to identify and carry out this research? Finally, what projects do we undertake to carry out this strategy and with what people and resources? The remainder of this whitepaper addresses the Intel Science and Technology Center for Secure Computing answers to these questions.

We have a dream

Rather than beginning with a list of projects, let's start with some "dreams" we have for our users. In each case, neither current products nor complete solutions pointed to by published research fully enable the dream. Solving key technical issues to fully enable the dream lies at the heart of the ISTC and the reminder of the whitepaper will chart a course to making the dream a reality.

Dream: Computing everywhere, any time

Everyone has a secure mobile device that they always carry. They are never out of touch and can get information, conduct business, get entertainment and communicate anytime, anywhere. People can use their devices for business (even important business matters demanding the utmost confidentiality) or personal matters. The business use is completely isolated from personal use and employers confidently allow their employees to conduct corporate activities using the device. Each user is also confident that, on personal matters, his phone neither steals nor leaks his information. Users no longer need to carry around a wallet; instead, they can shop and pay for everything using their phone. Despite the fact that the device can be used to pay for things and authenticate the user at airports, users don't worry about loss since their information is protected as well or better than it would be if it were in their locked houses. If their phone is lost or stolen, there is a convenient way to disable it, restore from a backup onto a new device, and be up and running with no loss of data or money. Users install many useful applications on their mobile devices but never worry about malware hurting important activities. They can get all their information anytime on their mobile device, which can access many data sources safely. The phone has a safety meter well calibrated to user expectations. The meter is seldom in any position but “green” for safe and it’s seldom wrong. If users are unsure about whether an action is safe, they can consult the safety meter, or obtain more details from a privacy and security dashboard.

Dream: Clouds from both sides

Corporations can take advantage of economical cloud infrastructure for their most important projects. They upload their software to data centers, which protect their data using security mechanisms whose safety has independently assessed and is constantly verified. Corporate information is only unencrypted inside the corporate software and is isolated from other data center tenants. Even data center “insiders” can’t get a tenant’s confidential information. Critical software operating in the data center has been verified for safe operation, and insurance companies offer low cost insurance for any corporate data loss when under the control of verified software. No claim has been filed for data loss and no data loss has been detected under these safe conditions despite repeated audits. Better still, different companies can cooperate on projects using these cloud data centers with better control and higher assurance than was formerly obtained by operating in a segregated space for each collaboration.

Everyone’s personal information can be saved in the cloud for safekeeping. This includes medical records, which, as a result, means authorized care providers can use this information anytime to provide treatment. Cloud software can use all kinds of user information to offer better services but that information is never available for any unauthorized use and no violations of the promised use restrictions has been observed despite constant attempts by misbehaving organizations with large resources.

Dream: Usable security enabled by advanced analytics

Today’s mobile devices are packed with sensors that are capable of gathering rich contextual information, such as location, wireless device signatures, ambient noise, and photographs. Using advanced security analytics, reliable contextual models can be constructed on these signals to assess the risk level of a user’s current environment. Then authentication mechanisms can scale up or down to match the current risk level, to decrease or eliminate the level of user interaction required to authenticate in some situations, improving usability without any effective impact on security.

To further reduce the burden on users, devices learn about each user’s privacy and security preferences by learning from users’ past responses to security and privacy warnings and prompts; this enables them to automatically respond to future security and privacy prompts in many cases, sparing the user from having to manage their own security explicitly. Information about security threats are aggregated and continuously analyzed in the cloud to determine when user involvement is truly needed; as a result, users are rarely interrupted or asked about security, and only when truly necessary.

For example, Bob is shopping in his favorite clothing store (e.g., determined by algorithms using GPS and Wi-Fi signals). He pulls his smart-phone out. As he moves the phone towards his face (e.g., detected by algorithms using accelerometer signals), the phone camera captures a series of images and performs facial recognition. Once the phone recognizes Bob, it unlocks the screen. Then Bob can open his mobile shopping app and use it to take a picture of a barcode on a jacket that caught his eye. The app informs him that the same jacket is available online in a variety of colors for less. He selects the color and size he wants and purchases it using his default credit card and shipping address. The store is a common destination for him, so Bob's password is not required to check out. Instead, he answers a multiple-choice question about his last purchase at the store; answering the question is a breeze compared to entering his password.

At home, Bob pulls out his tablet, which automatically recognizes him and unlocks the screen. He opens his mobile banking app and selects "Add a New Payee." The app takes his photo and compares the photo and his location to his existing profile. A moment later, Bob is able to add his cable provider to his list of payees without entering his long bank password.

Security Properties and Security Mechanisms

As noted, security and safety result from integrative properties of many elements of use: the environment, the activity, the actors and the technical mechanisms used to establish reliable and comprehensible properties that contribute to the desired outcome. New mechanisms and novel approaches can often be required in new situation. This is a core area for our research.

Paradigmatic properties that bear on security for user benefit and some of the mechanisms used to achieve them include:

Isolation: This property corresponds to the intuitive notion that activities can be separated so they do not adversely affect each other. In the physical world, enclosures like separate structures, locked rooms, binds and physical separation provide this property. In computer technology, separated computers on "air gapped" networks, CPU memory isolation technology like virtual memory managed by an operating system and virtual isolations techniques like encryption, which serves to restrict access to information to unauthorized parties during transmission and use fill this need. Benefitting from a "connected world" requires relying on hardware, information, networks, and software from many parties. Unlike traditional physical isolation techniques, cyber isolation techniques sometimes become ineffective in the presence of subtle flaws and unmet assumptions. A major goal of the Center is to develop simple, comprehensive, flexible isolation that enables new services while preventing unseen and potentially catastrophic adverse effects in seemingly unrelated activities.

Confidentiality: This property involves limiting the ability of an unauthorized party to view or use private or "confidential" information. In a distributed environment, like the Internet, cryptographic mechanisms have been developed to protect information in transmission. Recently, these techniques have also been employed to protect information in use and activities in progress. Even more recently, cryptographic techniques have been employed to protect information and activities operating on the same computer hardware. It is important to protect confidentiality while allowing authorized sharing of information under policy control.

Integrity: This property protects information from being changed, modified or corrupted by error or unauthorized activities. In the physical world, simply restricting the production and use of critical information to authorized and trusted parties may provide adequate integrity protection; wax seals have been used to ensure letters have not been tampered with. Critical information like prescriptions, require integrity protection. In the rich, distributed cyber world, cryptographic mechanisms have been developed to allow users to verify that information comes for identified sources and have not been modified by others while preserving the ability to collect, store and analyze the information in many places.

Availability: This involves reliable, predictable access to a capability or service. Many people are familiar with *Denial of Service* attacks, which prevent or slow access to web sites. Several mechanisms can be employed to achieve availability including carefully controlling design and operation parameters or providing redundant operations.

Recoverability: This involves restoring operations after interruption or, in our case, a security failure. Recoverability is closely bound to the security properties that we seek to protect. In some cases, like denial of service, cheap effective redundant operations can be a completely satisfactory mechanism for achieving user goals. Detecting data tampering and being able to resort to "known good data" is another example. Breaches in confidentiality or isolation are often much more difficult to recover from because of the speed of dissemination.

Auditability: Banks often make errors but employ procedures that document elements of transactions that both enables remedial action when possible, and accurate assessment of loss when remedies are incomplete.

Situational awareness: This is a critical aspect of security and is often poorly understood. It is virtually impossible for a user to adopt a security strategy incorporating mechanisms or procedures without a clear idea of the risks and rewards. These include the number and capabilities of adversaries trying to attack users, the effectiveness of preventive measures, the likely damages resulting from a breach, the potential recoverability after a breach and a user's ability to detect breaches when they happen. Computer security researchers often envy the analogous physical world security models in which theft is obvious (my car is gone), recoverability is intuitive (buy a new car), and cost is easy to gauge. Not so in the cyber world: connected adversaries are powerful and numerous, it is very difficult to determine accurately if an attack succeeded or who mounted it, the damage is often difficult to assess and most users (even professional computer scientists) have trouble determining the value and effectiveness of some technical protective mechanisms. Security analytics over large data sources is a new and important technique to assess these characteristics and hopefully even predict realistic outcomes.

Authorization: This property involves the specification of access rights to resources by users and enforcement techniques that restrict access to the specification. In order for this to be effective, users must be able to understand policy choices and elect them simply and reliably. Application permission systems on mobile phones, access control systems on files and records are mechanisms to achieve this control.

Authentication: Determining the person or thing requesting access to a resource or the person or thing responsible for discharging a security obligation (like providing confidentiality for a transmission) is a critical prerequisite for most security decisions. In the physical world, passports and drivers' licenses serve this purpose as do manufacturers marks ("Yale locks"). Passwords are a well used and often misused mechanism to provide authentication for people. Cryptographic techniques have emerged that perform user authentication more reliably and they have been extended to authenticate devices and software, which, after all, are ultimately responsible for computer security.

Usability: In many cases, users are asked to make security decisions: With whom can my data be shared? Can I install this app? Should I click on this link? Yet even when choices are obvious, users are asked "out of the blue" and in a manner that defies effective use. It's hard to take back bad choices and almost impossible to figure out what damage a bad choice caused. Even when choices clearly need to be made, they are presented badly so users fail to make the choices they actually want to. User error and misconfiguration remain one of the leading sources of security compromises, yet this subject has seen relatively little study among the security community. Thus, usable security is a key focus area for the Center.

Assurance: This property involves the ability to determine whether something performs its function in the expected manner. In the cyber world, we rely on digital hardware and software to insure isolation confidentiality, integrity, and auditing properties, enforce stated safe sharing policy, as well as be free of exploitable flaws that allow adversaries to subvert our purpose. Cyber security people look with envy on much of the physical infrastructure when considering assurance: People can tell if their houses protect them from the rain or their windows prevent outsiders from seeing inside their houses, but it is much harder to tell whether your computer will protect you adequately from attack. Mostly, people can tell if physical devices work: you know if your car starts, accelerates, turns and brakes (though this changes when computer systems control your car, as has been recently demonstrated). This is not at all clear in the cyber world. Major flaws in commercial software provide numerous examples of unexpected (and for users, undetected and, in many cases, undetectable) breaches in confidentiality, integrity, isolation and authentication failures, as well as policy enforcement failures due to software flaws. Microsoft, for example, has released many, many patches over the last ten years and continues to do so on a regular basis. Many of these flaws allow an adversary complete control over a user's machine and complete access to view, store and modify all user information; users can become vulnerable simply by connecting to the internet and using their software as intended. A major goal of the center is to provide reliable assurance by verifying that artifacts carry out their intended function without security flaws.

Research Themes

The foregoing landscape points out the glaring importance of some critical problems that must be solved to make our dreams possible:

1. Assure that the design of well scoped and well-understood computing activities operate as expected with few or no critical flaws.
2. Ensure that different computing activities are isolated so errors, flaws or limitations in functions implementing activities cannot compromise other activities operating on the same physical device.
3. Develop agile security mechanisms that can operate safely even in environments (like networks), that for reasons of economy cannot be made intrinsically safe.
4. Develop user-facing components that allow users to specify protection requirements for their activities and information, and understand effectiveness, potential losses and potential gains of that policy.
5. Develop tools and techniques that determine underlying risk and hidden benefits by analyzing the vast sensor, social media and unencrypted communications data. Ultimately, develop a faithful model that predicts the safety of activities.
6. Extend security mechanisms to ensure that operations are safe even on remotely sited equipment and that data is protected (i.e., used only in accordance with user policy) wherever it is used while not limiting expanded authorized access.

For each of these, the Center has identified critical technology, as well as foundational artifacts whose widespread adoption would provide satisfying and tractable solutions to these critical problems. These include:

1. Tools to verify security properties of well encapsulated systems: Static and dynamic analysis as well as information flow based analysis to determine that

hardware and software components carry out their intended function while maintaining critical security properties like isolation, confidentiality and integrity.

2. Simple reliable hardware rooted functions that allow isolation, authentication and even remote verification of all security critical components for encapsulated activities. This provides a foundation for protecting confidentiality and integrity as well as a practical hosting environment for encapsulated systems thus allowing devices, networks and services to participate in multiple activities without interference.
3. Tools and techniques allowing users to “say what they mean” when setting security rules and understand the impact of those rules, not just theoretically, but in a manner that enables them to make safe, informed choices.
4. Artifacts, tools and techniques to ensure the transparent and reliable enforcement of security properties so that “what you get is what you expect.”
5. Tools and techniques to provide situational awareness with respect to attacks as well as technology to learn new and valuable information to discover and incorporate novel protection mechanisms.

Research Focus Areas and some current projects

Mobile (User transparency, authentication, isolation)

The mobile effort is working to establish a secure foundation for mobile devices, and particularly for mobile apps.

Mobile apps have played a critical role in the success of smart-phones and tablets. They enable rich customization of these devices, functionality for the long tail of users, and a thriving economic marketplace for innovation. Our research seeks to ensure that apps remain safe for users.

We believe that strong security helps facilitate a successful economy for mobile apps. When users feel comfortable and safe installing apps, they are more likely to install apps from developers they have never heard of -- which in turn enables new entrants and new innovation and removes barriers to entry for app developers. In this way, security serves as an enabler for robust competition and a thriving mobile ecosystem.

One of our guiding goals is to help mobile platforms avoid the sort of malware problems that currently plague the desktop world. At present, desktop machines suffer from malware and greyware, as a result of architectural shortcomings in their security model. The result is frequent security compromises, anti-virus software that degrades system performance, and ultimately, frustration and inconvenience for end users. Unfortunately, the burden of legacy applications makes malware on the desktop very challenging to fight at this point. Fundamental aspects of the desktop architecture were designed and became entrenched in a less-risky era where attacks were rare. And, once a vulnerable architecture becomes entrenched, legacy constraints make it very difficult to make the kind of bold, radical changes that would be needed to address the fundamental security shortcomings in the architecture.

There is a grand opportunity to establish solid security foundations for mobile platforms now, before mobile systems become locked in by legacy constraints. The mobile space provides new opportunities to protect users. App stores provide a foundation for security services, such as analysis of apps to detect and respond to malware. The new platform provides an opportunity for new security models that respond to the threats users actually face.

Our work builds on the success of existing app markets and mobile security models. App markets arguably represent one of the most promising directions for security in the past few years. By providing a central location for users to find and install applications, they give end users a way to evaluate apps (e.g., through ratings and reviews from other users) and potentially to understand the risks associated with the app (e.g., through security permission models enforced by the mobile platform). App markets give market operators an opportunity to review apps in advance for malware, a way to monitor app installations, and, in situations where someone discovers a malicious app in the wild, an effective way to quickly disable that app and safeguard all affected customers. And, app markets create barriers to entry for malware writers. Since many end users will likely be more reluctant to install apps that have few positive reviews and few installations by other users, would-be malware authors are forced to develop production-quality apps if they want enough victims. This increases costs for attackers and drives malware economics in a direction that improves overall system security.

To achieve this vision, we must solve a number of technical challenges.

First, scale poses a major challenge. We must devise solutions that scale to the hundreds of thousands of diverse apps found in app markets, without restricting or limiting application functionality. And, as hundreds of new apps are submitted each day, automated analysis and proper incentive design is essential to scale to this magnitude.

Early research from the center has shown that many users are not able to understand current app permission systems. App permission systems must become more usable. Towards this end, we are studying how to design app permission systems that are effective yet non-intrusive. We are studying how to give users greater visibility into what their apps are doing with their data, and greater control over their apps. There are also opportunities for mobile hardware to provide innovative new capabilities (e.g., taint tracking) to support these goals.

We also need a better understanding of how to strengthen the feedback loops in app markets to better support security. Ideally, risky, malicious, or insecure apps would receive poor ratings, warning other users off from them and thus providing an incentive to developers to build apps that benefit end users and not put their security at risk. This requires us to better understand the entire ecosystem and develop feedback mechanisms that contribute to the health of the ecosystem and disincentivize malicious behavior. Perhaps apps should come with a product label that identifies their security and privacy attributes? Perhaps we can make better use of app reviews and ratings to protect users? Opportunities abound.

We are also studying how to make security on mobile platforms more usable and useful for end users, and how to safely integrate mobile and cloud. Currently, authentication on mobile platforms can be a pain: entering passwords is an annoyance, and more-convenient methods often offer weaker security. Also, there are widespread concerns over theft or loss of mobile devices: many end users have either experienced this personally, or know someone who has. Thus, one important challenge is to develop technology to mitigate these threats and ensure that your data will remain safe (and backed up) even if your mobile device is stolen or lost.

We expect that innovations in this space will have impact beyond the mobile world. Indeed, several desktop operating systems are already starting to partly adopt app markets and other security ideas from the mobile world. The time is ripe for research in mobile security to make computing more secure for end users across a broad range of computing devices.

The Secure Thin Intermediation Layer (STIL)

Among the security properties required to make our dream a reality are reliable isolation, authentication based on code identity (e.g.-attestation), and assurance, based on careful security property verification. Confidentiality and integrity mechanisms use these properties as a foundation. On computers running many different applications with different requirements this has been difficult to achieve, historically. A *Secure Thin Intermediation Layer (STIL)*, typically implemented as a hypervisor or security kernel, makes this possible and practical. The STIL enables well circumscribed, highly secure activities to run on machines running unrelated (and even adversarial) programs in clouds, on client PCs, and on mobile devices.

The STIL effort seeks to establish this foundation for safe computing in a way that is simultaneously *deployable*, *trustworthy*, and *flexible*. To be deployable, a STIL must be able to work with existing software providing at least sandboxing; users will not benefit from STIL capabilities without this requirement, given the wealth of deployed, complex, and insecure systems and application software. Taking advantage of STIL capabilities to provide additional capabilities to fully protect users, may well require modification of operating systems and applications To be trustworthy, a STIL should be provably secure, using rigorous design, validation, and verification technology; a STIL can and should be much smaller than entire operating systems, STIL verification is achievable within the center's mandate. Finally, to be flexible, a STIL must allow all the modes of interaction the users of computing systems enjoy today; it has to accommodate current applications, but also adapt to new modes of interaction that are sure to come.

Towards developing a STIL, the center is pursuing three major research goals:

1. We seek to understand what primitives belong in a STIL and what primitives are better implemented by higher software layers, such as operating systems. This effort must balance the need for a slim, simple trusted STIL with the expressive needs of security software above. We consider primitives such as isolation, software and user identification, execution monitoring, and trusted paths between input/output devices and users.
2. We are pursuing verification and validation techniques particularly targeting low-level software, to bring us closer to a provably secure STIL. The complexity of low-level software, which must manipulate hardware efficiently with minimal overhead, creates significant technical challenges for this effort. What is more, a STIL *may* typically need to emulate hardware; efficiency requirements increase the complexity of emulation, e.g., through the use of dynamic binary translation. Validating and verifying emulators of hardware is still an open, daunting research problem whose solution can positively impact the safety and trustworthiness of STILs.
3. We are undertaking the design and implementation of a STIL – an isolating multi-guest hypervisor – that can be used to validate our research, as well as innovative user-facing applications that demonstrate the value of the STIL.

Center researchers have launched a number of research projects targeting all of these research goals. We briefly mention a few below.

Hypervisor Attestation

This project places an integrity measurement of the hypervisor (i.e., the cryptographic hash of the hypervisor's booted image) in hardware, so that it cannot be undetectably corrupted by software. In connection with this measurement, hardware can protect secrets (usually keys) for the hypervisor, which, in turn, can provide similar services for its hosted operating systems. Finally, this measurement, coupled with cryptographic functions in the hardware can be used to remotely *authenticate* that an undamaged, properly verified hypervisor is running before confidential data is disclosed or processed by the hypervisor or designated guests. The ability to use this flexible "hardware root" for safe computing is a major theme of the ISTC.

Safe Remote Computing

Several of the center's projects pursue safe remote computing rooted in hardware. For example, they allow cloud "tenants" to run software in a cloud data center in a manner that ensures the integrity of operation of the software and the ability for the cloud-hosted software to encrypt all output (and decrypt all required input) under safe key management, thus preventing even cloud data-center “insiders” (under modest operating requirements) from corrupting or viewing confidential data. A related project uses these techniques to ensure that user data remains confidential and used only in accordance with user-specified rules anywhere that data is stored and processed. Finally, another project builds a minimal “cloud terminal” as a trustworthy user interface from an otherwise untrusted client OS to a sensitive application, such as a banking portal.

Verifiable Execution and Resource Accounting

STIL researchers are studying how to obtain a verifiable account of how a program executed at a remote platform (e.g., at an Infrastructure-as-a-Service provider’s site). The project plans to use a combination of hardware traps (collected by the STIL) and features such as instruction-tracing to track program events including context switches, I/O requests to external devices, memory use, and control-flow transfers. Such tracked events can then be used for runtime monitoring at a higher software layer. One such use considered is the verifiable accounting of resource use by a customer’s program at a provider’s platform, to ensure that pay-per-use invoicing is accurate. The project also considers other monitoring uses such as attested execution (an accurate log of exactly what a program did), control-flow integrity (protection from high-jacked jumps or returns by malware), and protection from kernel-level malware.

Verification Techniques for System Software

STILs typically rely on large data structures, such as caches, page tables, software Translation Look-aside Buffers, etc., to perform their duties faster. Center researchers are working on verifying formally that such structures are correctly used by a STIL. This problem is challenging because most automated verification methods (e.g., model checking) cannot handle the sizes of such data structures. This project uses safe abstraction techniques to produce a smaller model of how the STIL uses such large structures; on which automated verification can be feasibly performed. Those techniques are sound: they guarantee if the small abstract model is found secure, then so will the original model characterizing the entire STIL. A related project aims to validate that a model of a how the STIL manages its data structure accurately characterizes what the STIL software does in practice.

Finally, STIL researchers are compiling a compendium of Grand Verification Challenges for system software, by identifying the main security primitives of STILs, understanding their implications for software verification, and the opportunities for innovation in verification techniques motivated by STIL design and characteristics.

Collaborative research

STIL research is essential to clouds, mobile devices, embedded devices, infrastructure components, and PC’s; consequently, the STIL agenda cuts across many of the ISTCs. For example, use of STIL technology provides high assurance for applications running in shared cloud infrastructures. Client and cloud applications can use the trust management (i.e., attestation) enabled by “safe STILs” to extend the security properties to distributed application elements collaborating to provide a shared application-wide security goal. Privacy-preserving analytics and “big-data” computation using diverse and sensitive data sets are also enabled by this technology to support data-privacy objectives. In pervasive-computing use cases, the ability to extend trust boundaries is indispensable to ensure worry-free adoption.

Security analytics (situational awareness, recovery)

The world of security changes rapidly. New threats emerge on a rapid time scale. How can we keep up? Advances in “big data” and machine learning provide an exciting opportunity to help manage and monitor security-critical systems and adapt to new threats as they emerge, through new methods for security analytics.

With the advance of large-scale infrastructures for distributed sensing and computing (e.g., honeypots, malware collectors, smart-phones, crowd-sourcing platforms, etc.), abundant data can be collected from numerous (widely distributed) areas and devices. Deriving value from these large-scale data sets requires both powerful and reliable analytic tools. These tools pose both opportunities and challenges for security analytics.

Security analytics tools, many of which are based on machine learning and data mining techniques, are used to tackle the large data sets associated with a broad set of computer security problems, for example, classifying attacks using clustering algorithms, predicting attacks by using learning algorithms, detecting attacks through artificial intelligence, etc. Given these voluminous data sets, successful machine learning and data analytics for security applications such as anomaly detection and attack classification requires simultaneously identifying relevant features and constructing predictive models. However, in today’s systems, features are mainly manually selected by domain experts, which usually results in approaches that scales poorly as new systems are developed and new usage cases arise. As such, we need novel techniques for automated feature extraction and model construction from programs and large-scale data sets.

In addition, traditional machine learning methods have two flaws when operating in the presence of adversaries (adversarial environments) — the algorithms are based on statistical assumptions about the distribution of the input data, and they rely on training data derived from the input data to construct models for analyses. Adversaries may exploit these flaws to disrupt analytics, cause analytics to fail, or engage in malicious activity that fails to be detected.

In our current work, we are: (a) working actively with Intel/McAfee to identify applications and large datasets that we can analyze and perform security analytics upon, in addition to applying our techniques to the significant amounts of measurement/sensor data that we expect the STIL and secure mobile devices research efforts will generate; (b) developing tools to reliably and automatically extract features and build models for security applications by combining techniques from machine learning and program analysis; (c) developing tools to improve the user experience in interacting with computer devices for security operations; and (d) developing counter-measures to adversarial manipulation of metrics. A couple of sample projects are described in more detail in follows.

Secure Authentication with Face and Eyes

Many mobile devices now offer face authentication as an alternative to passwords or PINs for login. However, the current generation of face authentication systems is easily spoofed. We demonstrate that relatively low-resolution photographs can be used to successfully attack some commercially available systems, and also show that even systems that employ liveness detection can be spoofed by other easily-generated multimedia documents, including videos, slide shows and animated 3D models.

To overcome the discovered vulnerabilities, we are developing SAFE (Secure Authentication on with Face and Eyes), a system for logging into mobile devices. SAFE is designed to authenticate physically present users, applying face recognition for identification and gaze tracking for input of a secret. We assess the practicality of gaze motion as a means for secret input in a user study and examine the tradeoffs between uncertainty in the accuracy of the gaze tracker, usability, and the security of our system.

Robust Detection of Comment Spam

We aim to design a method for blog comment spam detection using the assumption that spam is any kind of uninformative content. To measure the ``informativeness'' of a set of blog comments, we construct a language and tokenization independent metric which we call content complexity, and leverage this metric to create a small set of features well-adjusted to comment spam detection by computing the content complexity over groupings of messages sharing the same author, the same sender IP, the same included links, etc. Using these features, we derive a simple mislabeling tolerant logistic regression algorithm and train a classification model based on expectation-maximization method. We evaluate our method against an exact set of tens of millions of comments collected over a four months period and containing a variety of websites, including blogs and news sites, we show that our method can operate at an arbitrary high precision level, and that it significantly dominates, both in terms of precision and recall, the original labeling.

Large-Scale Efficient System for Detecting Spam URLs

Existing research has shown promising results in using machine learning based approaches for detecting spam URLs based on features related to the HTML, page content, JavaScripts, etc. However, most existing approaches rely on human expert to identify (hundreds) of useful features, and the feature set has to continuously evolve as the malicious URLs evolve, which demands large non-stopping effort from human experts. We would like to develop an automated process to identify features without human experts by combining techniques from both program analysis and machine learning.

In addition, expert effort is also required for collecting and labeling training data, handling edge-cases, and providing unbiased evaluation metrics. To optimize expert effort in building and evolving the prediction model, we will explore machine learning techniques (e.g., active learning, ensemble method and stratified sampling) to identify a minimal subset of most effective data to for expert to collect/label for improving and adapting the model, and to develop new machine learning algorithms that can tolerate label noise with little performance degradation.

Research Style

The Intel Science and Technology Center for Secure Computing follows an open research model. We publish our results as well as the artifacts we develop including software under open source licenses and make a determined effort not to erect intellectual property tolls or barriers so that others can build on our work.

The ISTC is headquartered at UC Berkeley, and also encompasses faculty and students from four other academic institutions: Carnegie Mellon University, Drexel University, Duke University, and University of Illinois at Urbana Champaign. We also collaborate with other academic and industrial researchers including Intel Corporation when that work does not conflict with our open IP policies.

The Center holds two retreats each year at which Intel Corporation collaborators and sponsors and ISTC members meet, review results and opportunities and start new project collaboration opportunities. Intel Labs builds on our work to develop new capabilities for Intel products.

People

The Intel Science and Technology Center operates under the direction of two co-Principal Investigators, David Wagner from the Computer Science Department at UC, Berkeley and John Manferdelli of Intel Labs. In additions to John, three Intel researchers, Petros Maniatis, Ling Huang and Vyas Sekar work in the Computer Science Building at UC, Berkeley; they often teach classes and advise students in connection with ISTC research. Berkeley faculty members in the center include

David Culler, Anthony Joseph (who is Associate Director of the Center), Doug Tygar, Dawn Song, and Sanjit Seshia. Four “spoke schools” form a critical part of the Center. These are Duke University, where Landon Cox is the principal faculty participant, Drexel University, where Rachel Greenstadt is the principal faculty participant, Carnegie Mellon, where Adrian Perrig is the Principal CMU faculty participant and University of Illinois at Urbana Champaign where Sam King is the principal faculty participant.

The center supports twenty graduate students and post-doctoral staff at Berkeley and the Spoke schools. Aside from the four "embedded" Intel Labs researchers, other Intel Labs members also collaborate on research projects. Researchers at other institutions like Cornell and Google, while not funded by the center, work on research projects together with Center members on projects of common interest. Finally, researchers at several other Intel-funded institutions (The ISTC for Cloud Computing at CMU, the ISTC for Pervasive Computing at University of Washington, the ISTC for Embedded Computing at CMU and the ISRC for Security at Darmstadt) also work with Center members. The ISTC is an open community that welcomes other sponsors under the open IP model.

Making the dream come true

The mission of the Intel Science and Technology Center for Secure Computing is to carry put research to explore and develop new ideas to help make the vision of secure computing a reality.

We've made the point that what makes security so important is the interdependent operation of devices often connected by public network. This distributed computing model confers tremendous benefits and is at the heart of new capabilities that increase the value of computing for everyone. Unfortunately, this self same distributed model is the source of our greatest security challenges, challenges not effectively solved by earlier security techniques that evolved in the world of private data centers and time sharing.

The particular areas that raise alarm are: malware (downloaded in the connected world), authentication (you don't see the software or people upon whom you rely) and breach of confidentiality and integrity (the data upon which rely comes from more varied sources and used at more varied locations) and lack of insight (attacks happen silently, complex policy decisions () and often tremendous damage is done in milliseconds with no clearly observable signs). To achieve the security mission for users benefit, the center has elected to on key unsolved problems: authenticated and verifiable isolation so activities can be safe on local machines and in the cloud even when operating on devices which run malware or co-host adversaries, better analytics to improve visibility into risks and losses, useable policy models for users do that they can effectively dictate policy and understand consequent enforcement for their valuable data and safe devices, like mobile phones, to provide flexibility and access to information services anywhere, anytime.

Fortunately, we think we can solve the basic security challenges to make our dreams real and we've selected research topics that are critical to doing just that. By combining this research with existing security technology, we can work with Intel and others to build the things to make this happen.

Follow us

Below is a list of papers from the first nine months of center operations. You can follow the center work at http://scrub.cs.berkeley.edu where papers and software releases ate posted.

Software will be posted on http://scrub.cs.berkeley.edu/software/ when released.