Freezing More Than Bits: Chilling Effects of the OLPC XO Security Model

Meredith L. Patterson
University of Iowa

Len Sassaman
Katholieke Universiteit Leuven

David Chaum
Katholieke Universiteit Leuven

Abstract

In this paper, we discuss *Bitfrost*, the security model developed by the One Laptop Per Child project for its XO laptop computers. Bitfrost implements a number of security measures intended primarily to deter theft and malware, but which also introduce severe threats to data security and individual privacy. We describe several of the technical provisions in Bitfrost, outline the risks they enable, and consider their legal ramifications and the psychological impact posed for children and society.

1 Introduction

Since its announcement in 2005 [13], the OLPC XO laptop computer has been hailed as a revolutionary innovation in the quest to bring computer literacy to the majority of the world's population. The small, sturdy laptop is extremely inexpensive, consumes very little power (and can be charged with a hand crank or foot pedal), has no failure-prone moving parts, provides wireless mesh networking, includes a built-in video camera and microphone, and features a novel graphical user interface (known as *Sugar*) which is intended to "turn the laptop into a fun, easy-to-use, social experience that promotes sharing and learning" [18]. To date, the governments of Argentina, Brazil, Libya, Nigeria, Peru, Rwanda, Thailand and Uruguay have agreed to purchase XOs for their schoolchildren; it is estimated that between 5 and 10 million XOs will be distributed in 2008 [15]. The first deployments of XOs have already begun in Mongolia [17] and Uruguay [12].

Due to concerns about theft, the XO design team has taken measures to render the laptop a less attractive target for illicit resale. Most components are soldered directly to the motherboard, to discourage parting out the machines. The XO also implements a software and firmware security platform, dubbed *Bitfrost*, aimed at preventing theft, damage from malicious software, compromise of user privacy, and compromise by software which harms other network users (e.g. botnets or spam relays) [11]. Although these are noble goals, many of Bitfrost's provisions present much more dramatic risks to XO users than those the policy is intended to deter.

In this paper, we analyze the technical weaknesses of the Bitfrost security policy; enumerate the dangers which Bitfrost not only fails to prevent, but indeed actively *encourages*; and discuss the sociological ramifications of the human-computer interaction model which Bitfrost is poised to unleash on an unsuspecting user-base.

2 Technical Concerns

2.1 Principles, Goals, and Threat Model

The Bitfrost specification outlines four principles and five goals intended to guide the technical features of the platform: "Open design," "No lockdown," "No reading required," and "Unobtrusive security;" and "No user passwords," "No unencrypted authentication," "Out-of-the-box security," "Limited institutional public key infrastructure," and "No permanent data loss." These are laudable aspirations, particularly given that most of the XO's userbase will have had limited prior exposure to technology and many will be too young to read.

Bitfrost also establishes a five-point software threat model, intended to encompass the categories of " 'bad things' that software could do." It comprises:

- Damaging the machine;
- Compromising privacy;
- Damaging the user's data;
- Doing bad things to other people; and
- Impersonating the user.

These are quite reasonable threats to consider, and Bitfrost shows much promise in protecting its users from unauthorized abuses (intentional or accidental) from misbehavior of software applications.

The Bitfrost specification includes a lengthy list of hardware/firmware, kernel-space, and user-space policies and

`chroot` environments intended to prevent malicious software from accomplishing any of the above goals. The OLPC XO is designed such that it cannot be activated without complying with these policies, thus discouraging attempts to divert XOs away from the supply chain and onto the black market (a goal stated in section 3 of the specification). These measures will be costly and inconvenient to subvert.

However, many of Bitfrost's policies introduce more problems than they solve. We will examine several of these policies in detail, identifying areas where Bitfrost generates a scenario which diverges considerably from the rosy picture which its principles and goals paint.

2.2 A Peculiar Definition of "Open"

Although Bitfrost advocates open design, we note that the only available draft of the specification states that it is not the final version, and that a full technical specification is "being prepared" [11]. There is no indication that the specification has been submitted to any recognized standards body for approval, or even when a final draft will be made available.

Were Bitfrost still merely a proposal, this would not be such a cause for concern. However, 1000 XOs have already been deployed in Mongolia [17], and 8000 in Uruguay, with another 90,000 to be deployed in the next several months [12]. A *de facto* standard has thus been defined, in the form of the source code of the release builds of the operating system. Although the source code is publicly available, this alone does not constitute a standards specification. A true specification provides implementors with reference guidelines to use to verify the correctness of the implementation, and to ensure interoperability.[1] The lack of a formal specification bespeaks poor management practices, and leads us to question the quality of the implementation—if there is no standard, how is the platform to be tested?

In the remainder of this section, we discuss policies as they are described in the available documentation.

2.3 Digital Identity: the first-boot protocol

Each XO has a unique identifier tuple consisting of its serial number, SN, and a randomly-generated 32-byte identifier, $U\#$. When a country receives a shipment of deactivated XOs, it also receives the corresponding identifier tuples, and generates a unique *activation code* for each tuple. When the country sends XOs to a school, it also sends a USB key with the codes for each XO in a separate shipment; the school plugs this key into a server connected to a wireless network, which acts as an *activation server* for that batch of XOs. To activate an XO, a child powers it on within range of the activation server; the XO sends its identifier tuple to the server, which responds with the appropriate activation code, and the XO initiates its "first boot" process.

As the very first step of this process, the XO asks for the child's name and takes a digital photograph[2] of the child. It also generates an ECC keypair (without a passphrase; the key size is unspecified) and signs the name and photograph with this key. The resulting 8-tuple $\langle SN, U\#, N, P, ECC_{pub}, ECC_{private}, sig_N, sig_P \rangle$ forms the child's *digital identity*. It is immediately transmitted to the activation server (which serves as the primary backup server) and the country's central backup server.

Thus, the child is immediately linkable, by name and appearance, to the laptop he or she has been issued—and, more importantly, to a long-lived keypair which is now no longer under his or her sole control. We question the need for such invasive measures. The specification provides no rationale for storing the name and photograph, but presumably it is so that if a stolen laptop is recovered, its owner can strongly identify herself. Other biometric factors, e.g. voiceprints, might be a less privacy-invasive but equivalently strong means of satisfying this goal.

2.4 Data Security and Key Escrow

Recall that the Bitfrost specification explicitly lists "compromising the user's privacy" and "impersonating the user" as things that software running on the XO should not be able to achieve. However, without giving the user any other option[3], the XO transmits both halves of a keypair which is permanently associated with the user's identity to two separate entities, all before the user fully assumes control of the laptop! Bitfrost lists "limited institutional public key infrastructure" as one of its goals, but by default it establishes the most user-hostile form of key escrow [1]. The user has no control over the deposit, recovery or maintenance of her keypair; compromising a key store compromises all keys in it (since they have no passphrases), and the Bitfrost designers consider this an "acceptable risk" [11]. According to the P_DOCUMENT_BACKUP policy, this is motivated by a desire to avoid having to regenerate a child's digital identity if her XO is lost or destroyed. We question the importance of this goal, particularly given how unobtrusive the digital identity creation process is. The current structure requires key escrow for recovery of encrypted backups, but decoupling the data recovery process from the identity and authentication components would allow each problem to be addressed according to its specific requirements.

The P_DOCUMENT_BACKUP policy also allows any server advertising itself as a "backup service" to trigger automatic incremental backups of an XO's data. Although these backups are encrypted to the user's ECC key, this provides negligible protection against a skilled third party. Any individual who gains access to the key store (via "black-bag cryptanalysis" or "aluminum-briefcase cryptanalysis") can set up a backup service as a honeypot and compromise the private data of any XO in the "neighborhood".

2.5 Anonymity and Deniability

Thanks to Bitfrost's key escrow policy, it is trivial for anyone with access to an XO user's primary backup server to forge the user's signature on any document, with no way for the user to repudiate the signature. However, the threats Bitfrost poses against user anonymity are much farther-reaching than forged signatures.

The P_IDENT policy states that "all digital peer interactions or communication (e-mails, instant messages, and so forth) can be cryptographically signed to maintain integrity even as they're routed through potentially malicious peers on the mesh." Since the policy does not state the conditions under which traffic will or will not be signed, and the "unobtrusive security" goal emphasizes that "strong unobtrusive security" will occur "behind the scenes" unless it impacts usability—*not* privacy—we must assume that all outgoing traffic will be signed by default when possible. Since IP, TCP and UDP provide no mechanism for signing, this operation presumably takes place at the application layer, through overt message signing as described, or by signing the message body and embedding the signature in a header—the From request-header of HTTP [7] is an obvious candidate.

Signing, whether at the message or packet level, implies non-repudiability of all signed messages or packets. Ergo, it is impossible for XO users to use any form of anonymous communication with confidence.

The P_IDENT policy is thus a threat to many forms of speech which have been shielded by anonymity in the past: political speech, "whistleblowing" against corporate or governmental abuses of power, and religious speech, to name a few. (Granted, in the West, schoolchildren are not often in a position to expose corporate or governmental malfeasance—but in the Third World, corruption is often far more overt due to the belief of those in power that no one can do anything about it. The XO has great potential to empower the common citizen, but not if citizens cannot speak without fear of repercussion. In nations where it is not uncommon for schoolchildren to be drafted as soldiers, it is certainly possible for children to become whistleblowers.) The United Nations Universal Declaration of Human Rights protects not only the freedom of expression, but the right to privacy for member states' citizens [8]. Given that the OLPC project transacts with the national governments of UN member states, much more attention should have been paid to the security policy's effects on protected speech.

This policy additionally limits the utility of the XO by making it an unsuitable platform for networked voting systems in elections that require secret ballots. Nevertheless, S.T.I.R.M.E., an electronic voting project for the XO platform, is being developed [23]. If it is used beyond its current scope of classroom and open source project elections, S.T.I.R.M.E. could place users at risk or compromise election integrity due to the implications of the P_IDENT policy.

2.6 A Very Expensive Paperweight

XOs with the P_THEFT policy enabled must obtain a limited-duration lease—the specification suggests 21 days—from their home country's anti-theft server in order to remain activated. When an XO connects to the Internet, the P_THEFT daemon ("a privileged process that cannot be disabled or terminated even by the root user" [11]) "calls home" at most once per day to renew the lease. If an XO is reported stolen, the next time it attempts to renew its lease, the P_THEFT daemon shuts it down and returns it to a deactivated state. A new activation key is needed for the laptop to function again. If an XO's lease expires while it is not connected to the Internet, it likewise deactivates.

Leases can be renewed manually by means of a USB drive manually delivered to a school's activation server, but we question the utility of this approach in the event of natural disasters. Many of the target XO deployment locations are in remote, difficult-to-access areas which could be cut off from travel by earthquakes, floods or other catastrophes. If a school unexpectedly loses its Internet access for a long enough time, all its attached XOs will automatically deactivate, leaving students out of contact even after connectivity is restored (e.g., by repairing a broken satellite dish). This is at best inconvenient, and at worst, a serious hazard if people have come to rely on XOs as a primary means for long-distance communication.

More relevant from a security and privacy perspective, however, this policy is rife with potential for abuse. Combined with the anti-anonymity features of P_IDENT, P_THEFT is an extremely effective way of silencing specific individuals. Signed messages are linked to the XO they came from, so a government need only flag that XO as "stolen" in the anti-theft database in order to shut it off permanently. A country can also shut off all its XOs in one fell swoop by flagging them all, or simply shutting off the anti-theft server and waiting for all the leases to expire.

2.7 Replacement Firmware?

Children who become extremely proficient at working with the underlying components of their XO have the possibility of being granted "developer keys" that allow them to make modifications to the system, including potentially overwriting the existing firmware with their own software, or even their own operating system. The spec is unclear on how the precise mechanisms function in this case, but the existing spec proposes P_BIOS_COPY, a secondary BIOS containing an immutable copy of the primary BIOS firmware. This would allow the restoration of the original operating system and all of its controls, with no possibility of permanently disabling them. It is unclear under what circumstances this restoration can be invoked, or indeed what the limits of the the secondary BIOS's capabilities are.

3 Sociological Concerns

3.1 Human Rights and Chilling Effects

The privacy-eroding aspects of Bitfrost are of particular concern when one examines the human-rights records of the countries enrolled in the OLPC program. In Libya, criticizing the government is grounds for arrest and torture [2]. In Nigeria, citizens who speak out against government corruption face threats and physical violence, which has deterred civil rights groups from speaking up [9]. In Thailand, political activists have reported illegal surveillance by the military junta which took power in September 2006, and which claims the right to detain citizens without charge [10].

According to the legal doctrine of *chilling effects*, an activity, e.g. criticizing a corrupt regime, "is chilled if people are deterred from participating in that activity", whether through punishment or merely the threat thereof [20]. Bitfrost's design may not *intend* to facilitate surveillance on children, but as we have shown, it certainly does so. Combined with the powers the P_THEFT policy provides, it is easy to envision a scenario where a child blogs or e-mails a document which the government wants to quash, it is traced back to the child, and the child's XO is suddenly reported "stolen" and deactivated. Fear of a similar punishment would certainly chill controversial speech on the part of other XO users.

3.2 Habituation and Indoctrination

Founder Nicholas Negroponte says of OLPC, "It's an education project, not a laptop project." Taking a cue from the field of educational psychology, we examine the lessons that Bitfrost is likely to impart to XO users.

The XO's target audience is children between the ages of 6 and 12 [16]. In Piaget's theory of cognitive development [19], this corresponds to the *concrete operational* stage, when children acquire logical reasoning abilities and use them to form automatic working models of the world, or *schemas*. Erikson's theory of psychosocial development associates this age group with the *psychosocial crisis* of "industry vs. inferiority," wherein children are eager to learn but afraid of failure and punishment [6]. This is a pivotal stage of emotional growth, and the schemas children form during this timeframe persist for years. Traumatic events—particularly ones indirectly connected to a cause, such as being punished for "unapproved" speech by having one's laptop suddenly deactivate seemingly on its own—may have dramatic and long-lived negative effects on a child's view of the world and her place in it [4]. Even seemingly innocuous events can have an insidious effect on schema formation; children who grow up learning that handing over their identity to a remote authority is the "price" of Internet access may internalize giving up their right to privacy as a commonplace, expected event.[4]

Elliot Turiel's *domain theory* distinguishes between *moral values*, which are universalizable beliefs founded in concepts of justice, rights, and welfare; and *social conventions*, context-dependent standards of behavior tied to the social system [22]. Bitfrost's policies enforce a set of social conventions starkly at odds with those of the broader Internet. On the Bitfrost Internet, children may learn to view controversial speech as dangerous due to the risk of punishment, rather than a fact of life. This puts them at risk of failing to develop an autonomous sense of social responsibility, since the imposed social convention makes it difficult for children to identify the moral values which underpin responsible Internet citizenship [24]; given the conditioning they are subject to, they may come to advocate censorship and anti-anonymity policies which negatively affect the rest of the world, as well.

The Internet's predecessor, DARPAnet, was designed to be robust in the event of physical damage, providing flexible rerouting if a previous path becomes unusable. This architecture has given rise to John Gilmore's famous remark, "The Internet perceives censorship as damage and routes around it." However, if the P_IDENT policy extends to signing of all traffic, or if the P_DOCUMENT_BACKUP policy extends to archiving students' browsing histories (which can then be examined for "forbidden" content), this is no longer an option—a child's Internet access can simply be cut off at the source. This is a profoundly depersonalizing act, and one which threatens a child's sense of individuality and personal agency [14]. People have a right to expect that what they read, write and create, their correspondence and recreation, are a matter of personal choice. Subjecting children to constant surveillance damages their ability to establish personal boundaries and identify as an individual within a society; and yet the Bitfrost model opens the door to precisely that.

3.3 Imagined Communities

The XO is designed for use focused around local schools. Thus, the designers should be aware of the threats that users may face due to the misperception that their data is only accessible locally, or that they are only speaking to individuals within their own communities. For an in-depth look at the impact of "imagined communities", those that appear restricted to a given boundary but are in fact open to the Internet as a whole, we refer to Acquisti and Gross [3]. While this work focuses on the impact that social network sites with imagined communities have upon their users' behavior, the principle can be extended to any scenario where an imagined community may be perceived by the user.

Further research into the impact the XO local network and Internet interaction has upon the users of these systems will be needed once live deployments can be studied.

4 Conclusion and Future Work

Any security policy must be evaluated on its appropriateness and its efficacy: does it address threats users are likely to

face, and do its provisions actually mitigate threats? In this paper, we have examined several pieces of the Bitfrost security policy, and conclude that it suffers from an inappropriate threat model and an incomplete solution to the threats it outlines. Furthermore, several policies play a minimal role in the threat model, but expose children to threats which the Bitfrost model fails to include. The specification goes into great detail about what user-space code is not allowed to do, thus defining that threat model and protection bounds quite well. It does not give the hardware or operating system components the same level of scrutiny.

As there has been much work on privacy-preserving systems in recent years, it is our intuition that most, if not all, of the problematic aspects of Bitfrost can be eliminated by refining the specification to consider the dangers we have highlighted in this paper, while also considering the existing threat models. It would be ideal if we were able to work from a static specification, but we intend to experiment with replacement primitives for existing components in the draft spec to achieve the same security properties while eliminating the threats that the current methods introduce.

5 Acknowledgments

We would like to thank Lindsay Patterson for assistance with research into the psychological effects of traumatic events on children, and Wendy Seltzer for providing background material on the Chilling Effects Doctrine.

The work of Len Sassaman and David Chaum was supported in part by the Concerted Research Action (GOA) Ambiorics 2005/11 of the Flemish Government, by the IBBT (Flemish Government) and by the IAP Programme P6/26 BCRYPT of the Belgian State (Belgian Science Policy). Additional support was provided by the EU within the PRIME Project under contract IST-2002-507591.

References

[1] ABELSON, H., ANDERSON, R., BELLOVIN, S. M., BENALOH, J., BLAZE, M., DIFFIE, W., GILMORE, J., NEUMANN, P. G., RIVEST, R. L., SCHILLER, J. I., AND SCHNEIER, B. The risks of "key recovery", "key escrow", and "trusted third-party" encryption. *World Wide Web Journal 2*, 3 (1997).

[2] ABRAHAMS, F. Human rights in Libya. *New Statesman* (Jan 2008).

[3] ACQUISTI, A., AND GROSS, R. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *Proceedings of the 2006 Workshop on Privacy Enhancing Technologies* (2006).

[4] CASON, D. R., RESICK, P. A., AND WEAVER, T. L. Schematic integration of traumatic events. *Clinical Psychology Review*, 22 (2002), 131–153.

[5] DOCTOROW, C. Fingertip biometrics at Disney turnstiles: the mouse does its bit for the police state, Mar 2008. http://www.boingboing.net/2008/03/15/fingertip-biometrics.html.

[6] ERIKSON, E. *Childhood and Society*. Norton, 1950.

[7] FIELDING, R., GETTYS, J., MOGUL, J., FRYSTYK, H., MASINTER, L., LEACH, P., AND BERNERS-LEE, T. Hypertext Transfer Protocol – HTTP/1.1. Request for Comments: 2616, June 1999.

[8] Universal Declaration of Human Rights, Dec 1948.

[9] Nigeria: Investigate attacks on anticorruption campaigner, Mar 2007. http://hrw.org/english/docs/2007/03/07/nigeri15460.htm.

[10] Thailand: Military interference undermines upcoming elections, Dec 2007. http://hrw.org/english/docs/2007/12/20/thaila17631.htm.

[11] KRSTIC, I. The Bitfrost security platform, March 2007. http://wiki.laptop.org/go/OLPC_Bitfrost, retrieved 18 March 2008.

[12] KRSTIC, I. First OLPC deployment: now it's real, December 2007. http://radian.org/notebook/first-deployment.

[13] MARKOFF, J. At Davos, the Johnny Appleseed of the digital era shares his ambition to propagate a $100 laptop in developing countries. *The New York Times* (Jan 2005).

[14] NUCCI, L. The personal domain and formation of the individual. In *Values and Knowledge*, Reed, Turiel, and Brown, Eds. Lawrence Erlbaum, 1996.

[15] NYSTEDT, D. One million OLPC laptop orders confirmed. *IDG News Service* (Feb 2007).

[16] Core principles. http://wiki.laptop.org/go/Core_principles.

[17] OLPC Mongolia. http://wiki.laptop.org/go/OLPC_Mongolia.

[18] Sugar – OLPC. http://wiki.laptop.org/go/Sugar.

[19] PIAGET, J. *The Construction of Reality in the Child*. Basic Books, 1954.

[20] SCHAUER, F. Fear, risk, and the first amendment: Unraveling the "chilling effect". *Boston University Law Review*, 58 (1978).

[21] SCOTT, G. Guide for internet standards writers. Request for Comments: 2360, June 1998.

[22] TURIEL, E. *The Development of Social Knowledge: Morality and Convention*. Cambridge University Press, 1983.

[23] VERGARA, I. Secure transparent instant representative mesh elections, 2007. http://wiki.laptop.org/go/Stirme.

[24] WILLARD, N. Moral development in the information age. *Proceedings of the Families, Technology, and Education Conference* (1997).

Notes

[1]The Internet Engineering Task Force provides an excellent guideline for writing standards specifications in RFC 2360 [21]. While this is oriented toward the RFC series of documents published by the IETF, it can be used as a template for easily-readable and auditable standards published independently as well.

[2]While the OLPC design criteria calls for an LED on the activation circuit for the camera and microphone to discourage their use as surveillance devices, the developer models of the XO we have used lack this LED. It is unknown if the currently deployed units provide any visual status indicators for these hardware components.

[3]Possibly without notifying the user at all; the Bitfrost specification is silent on this issue.

[4]Privacy advocate Cory Doctorow relates a recent incident at Disney World, which has begun linking park visitors' tickets with a finger-geometry scan: "One morning at Epcot Center, as we offered our ID to the castmember at the turnstile and began to argue (again – they're very poorly trained on this point) that we could indeed opt to show ID instead of being printed, a small boy behind us chirped up, 'No, you have to be fingerprinted! Everybody has to be fingerprinted!'" [5]