November 29, 2020

Encrypting the Web, a “handoff”

Nick Doty
UC Berkeley, School of Information

Status of This Document *This is a case for exploration of the Handoffs model for analyzing tech policy and a section of a published dissertation: Enacting Privacy in Internet Standards.*

Encrypting the Web, a “handoff”

Cafe Confusion

You sit down at the little cafe on the corner, much like the cafe where I’m sitting and writing this story. It’s a lazy afternoon, you order a latte and while you’re waiting for it, you open your laptop and connect to the only open network: `FreeLittleCafeWiFi`. Why not open your email and see if your sister wrote you back? And before you even get to Yahoo! Mail, you think, actually, you should check Facebook and see if there are any new pictures of the nephews, and who’s coming to that party on Saturday night.

There’s a handful of other people in the cafe: a teenager probably from the high school across the street, a man with glasses tapping away on the next great American novel, a woman working on a presentation. Everyone is using a laptop.[1]

After a few minutes scrolling, scrolling, scrolling through the news feed, you notice that there’s a new post: it says it’s from you, but you certainly didn’t write a post of fart jokes. The teenager across the room snickers as she closes up her laptop and leaves the cafe.

[1] It’s 2010, say, and not quite everyone is using smartphones all the time yet.

Frustrated, you delete that post. What else did they see in your account? Hadn't you made sure to log in with the lock icon, and wasn't that lock icon supposed to protect you? Could people in the same cafe always see what you were doing online? Who else could do that? Shouldn't that be against the law? Shouldn't people know better? Did you do something wrong? *Who or what was really responsible?*

Handoffs

That uneasy question of responsibility arises from an unsettled combination of technical, legal and social processes. There is an implicit distribution among technical and non-technical means of assurance for a particular value of interest within a sociotechnical system. How you are able to communicate with friends and family over the Internet and whether those communications are secured from prying eyes and tampering depends both on the architecture of the Internet and the World Wide Web and on legal and normative protections of privacy and security.

In the example above, the mischievous teenager in the cafe might have used a small plugin called "Firesheep" that makes such eavesdropping and impersonation a straightforward point-and-click measure. The author of Firesheep, Eric Butler, makes his case for who is responsible: companies operating these websites *should have* implemented more widespread security with HTTPS for all connections and creating Firesheep was a way to expose this problem more clearly.[2] That position is well-argued, but the question of normative responsibility doesn't have singular answers; the Web could also be designed such that legal, rather than technical, protections were what disincentivized the attacker; or a technical system could prioritize accountability and auditing over protection against an attack. Or, as the argument is sometimes made from a certain reductionist perspective, the value of privacy might simply not exist in a certain setting, caveat emptor.[3]

Collaborators have defined a **handoff** as the transfer of a function or the responsibility for some value from one actor to another between two different configurations of a system (Deirdre K. Mulligan and Nissenbaum 2020). A movement towards encrypting the Web, more specifically to broadly increasing the fraction of Web traffic communicated over TLS-encrypted channels, is such a handoff, shifting the responsibility for security from that unsteady combination of factors to a deployed technical system of HTTPS in browsers and servers.

[2] The very directly titled presentation "Hey Web 2.0: Start protecting user privacy instead of pretending to."

[3] Perhaps, *caveat usus*, but I don't have the ideal translation for "user."

In this case, I detail the different actors that make up the socio-technical system that is the Web, its diversity of goals, and the handoff to a set of technical guarantees for providing the value of security in online browsing. We can see how the distribution of responsibility has changed, both a general shift in paradigm and a particular triggering event. How that handoff is being negotiated and implemented shows how values can be conceived, debated and enacted in a complex, distributed system.

System overview The Web as a **socio-technical system** is complex in both its makeup and function.[4] Billions of end users use web browsers on personal smartphones, laptops, smart televisions, desktop computers at their local library or Internet cafe. Web sites that those users visit are produced by newspapers, governments, corporations, non-profits, individual hobbyists; those sites are hosted on servers ranging from tiny low-powered devices sitting on a bookshelf to enormous server farms with distributed locations around the world. Interconnection between those end users and those servers typically happens over the Internet, itself an even larger system; communications typically hop from a local WiFi network, to a commercial ISP, to some series of backbone providers, to a commercial network provider, to a CDN or commercial server, and back again. Depending on routing protocols, peering arrangements, server distribution and local network infrastructure, those hops may cross many national boundaries and may take many different routes or all pass through a single undersea network cable.

A incomplete summary of relevant (human) **actors**:

- end users
- Web site developers
- browser developers
- ISP administrators
- advertising network executives

Or considering other *types of actors*, we might also identify key pieces of hardware or software: network switches, Web browsers, fiber-optic cables. Or institutional actors: diverse privacy laws in the US and EU, nation-state intelligence agencies, commercial security companies, organized crime syndicates.

[4] A similar overview of the Web as a socio-technical system opens the Do Not Track handoff case as well – these are written so that they can be read individually as discussion drafts for the handoffs model.

Diversity of goals Given the diverse users of the system and the broad spectrum of actors that compose it, the Web also has a wide range of goals or functions. Many people use the Web for personal communications: checking their email accounts, posting messages to their social network accounts, reading and writing blog posts. Commerce is a common set of functions: companies provide services for sale; people buy both digital and physical products; online advertising is widespread; media companies provide entertainment services. In part because the Web and the Internet can be used for quick personal communication, intelligence agencies also use the network for surveillance of different kinds, to review the messages of particular targets, to map social networks based on communications metadata, to detect new security threats.

For this case study, we will look at a single goal, or a related set of functions for which the Web could be designed: securing the communications between people and services.

Security is a broad, multi-faceted concept: consider the C-I-A triad (Clark and Wilson 1987) and Japanese *anshin* (Okumura, Shiraishi, and Iwata 2013). We identify security as a **value**. Confidential and integral communications could be considered the relevant **goal**, or the goal might simply be communicating with others and the intended **constraint** is for those communications to be widely available while being free from tampering or eavesdropping.

Paradigmatic changes

Implicit trust in the network Securing communications on the Web could potentially be accomplished through many different configurations of the socio-technical system. Historically (this is an overgeneralization, but stay with us), Web traffic was typically not encrypted between the endpoints. In order to facilitate online commerce – as users were concerned about entering their credit card numbers into such a new and less-understood system – transport-layer security standards were developed and many sites implemented that security for specific security-focused operations, like entering payment information or sending passwords for logging in to accounts.

This **paradigm** – *occasional security with user confirmation* – presumes trust in a range of network intermediaries. Assurance of the confidentiality of communications with your email provider, say, depended on the discretion of the ISP and other network backbone providers. Integrity of communications against modification by intermediaries was simply not provided as a technical matter; occasionally network providers would insert advertising or notifications for the subscriber.[5] Laws, regulations, norms and market forces could provide an incentive for those network intermediaries towards securing unencrypted communications against unwanted disclosure or troublesome tampering. Because those companies were typically regulated and within the jurisdiction of national governments, law enforcement or intelligence agencies had the technical and legal capability, at their discretion, to intercept any particular Internet communication. Relying on norms and legal backing, network operators, technical designers or expert users may have expected that such discretionary activity would be abnormal in the United States or other liberal democracies. Technical enforcement (transport-layer security, based on a PKI of certificate authorities) was most often used for explicitly sensitive data. Technical protection against downgrade attacks was limited or absent. Implicitly or explicitly, users had the responsibility to confirm through browser UI that a connection was secure before entering credit card numbers, passwords or sensitive information in order to obtain that technical support of confidentiality. Understanding error messages about the security of connections was challenging and users are faced with various seals and lock iconography with unclear implications (Sunshine et al. 2009).

Surveillance revelations as trigger To identify a singular trigger for the re-thinking and re-engineering of such a massive sociotechnical system, even limited to this particular function of secure communications, would be handwaving over a complex history. The position that the whole web should use HTTPS was common in certain communities before 2013, for various reasons related to privacy and security.

However, there are indications of a turning point in rhetoric and a substantial change in the momentum of action towards a new handoff configuration that can be related – in time and by explicitly-stated motivation – to the revelations in 2013 of widespread mass surveillance by the NSA and GCHQ.

Statements from engineers at the time indicate an acknowledgement of the previous handoff between state and technical actors, as well as the shift.[6]

A new paradigm: encrypted transport everywhere Driven by evidence of tampering with web traffic by ISPs and other intermediaries and widespread passive

[5] For example, Comcast has documented their notification system that inserts Javascript into web pages visited by the user (Chung et al. 2011).

[6] Dramatically: “we had a good thing / you messed it up [. . .] never again” (Thomson 2014).

surveillance by state actors, recommendations for Web security moved towards encrypted transport (HTTPS) being ubiquitous or expected for all (or most) kinds of Web usage. Rather than relying on the user to know when HTTPS was appropriate or necessary and manually confirm its use, servers were provided with the means to suggest or force usage of secure communications.[7] In this **paradigm** – *security for all Web traffic, driven by server and browser*, the user is out of the loop; Web communications are to have confidentiality, integrity and authentication by default, without user intervention, or even user understanding. In terms of threat modeling, the network is considered an attacker; widespread passive surveillance is directly addressed and not just for commercial activity but for personal information, various powerful Web capabilities, and for browsing activity in general; active downgrade attacks are mitigated; active, targeted man-in-the-middle attacks are made more observable.

Modalities of regulation during transition The shift described here – from occasional security to encryption everywhere – is remarkable in the breadth of re-engineering of technology and re-thinking of norms and practices in a large, diverse and not centrally-controlled group. To give an explanation of that transition might be to explain *why*, what motivated that change in paradigm, what upset the existing handoff and directed the community towards a different one. Identifying the *trigger* (above) is an attempt at such an explanation. Comparing the paradigms themselves and what actors are responsible for security in each is an explanation of *what* the handoff consists in. But another kind of explanation is to describe *how* a change is effected.

In using the handoff model, and as is common in analyses of tech policy, we can refer to different modes of action or different modalities of regulation. For example, from Lessig, we can refer to law, architecture, markets and norms as distinct modalities to regulate behavior, with distinctive properties (1999).

During this transitional time of negotiated re-engineering, the different groups of actors identified use different modalities of regulation; their activities are numerous and diverse. The actors and the modalities they try to use are perhaps not what we would initially assume.[8]

[7] In short: UIR, HSTS, the preload list.

[8] Whether this assumption is obvious or common I'm not sure, but I think we could typify government actors as using law, corporate actors as using market pressures, engineers as using architecture, advocates as focusing on norms.

Modalities of regulation interact substantially; there are rarely sharp boundaries. I attempt to group the actions employed during this transition by the modality of regulation that is predominant in each situation. In each case, the action is regulating in the sense that it influences the action of some other actor in our system, separate even from the actions that regulate the ongoing activity within our new or old handoff configurations.

Market Centralization in the technology field means that many of the companies that compete in one market also play a role in others. Microsoft famously produces and sells operating systems (Windows), and is also a significant browser developer (Internet Explorer) and operates a search engine (Bing), web sites (MSN) and online advertising. That multiplicity means that a browser vendor might use an alternative corporate role to influence a development of the Web. Google announced (Bahajji and Illyes 2014) that sites served over HTTPS would receive a boost in search results rankings.[9] Given the commercial importance of appearing high on a Google search results page (see: the SEO market), Web site operators had a new incentive to adopt HTTPS, even if it might incur the cost of purchasing certificates or upgrading hardware and software.

Corporate actors weren't the only ones to identify market incentives as important to this engineering change. The Let's Encrypt project was a collaboration between key companies (browser vendors, CDNs) and non-profit advocates (EFF) to establish a new certificate authority (CA), in many ways in direct competition with commercial players. Most significantly, Let's Encrypt provides the certificates necessary for authenticated HTTPS web sites at no cost. Where previously a small web site developer might have had to pay on the order of $10 a year to purchase and renew a certificate, Let's Encrypt made the process free and mostly automated. This was no doubt an application of direct economic incentives, but it also played a substantial rhetorical role in the larger process of convincing reluctant developers to embrace adoption of a new technology.

Architecture One debate that illustrates the particular uses of architectural features was the proposal to add "opportunistic encryption" to the HTTP standard. Different proposals might have operationalized that differently, but the suggestion was for servers and browsers to negotiated an unauthenticated encrypted channel

[9] If it weren't so beneficial for end users, we might expect that to fall under anti-trust scrutiny, as when the Department of Justice investigated Microsoft for using its OS monopoly to influence the Web browser market.

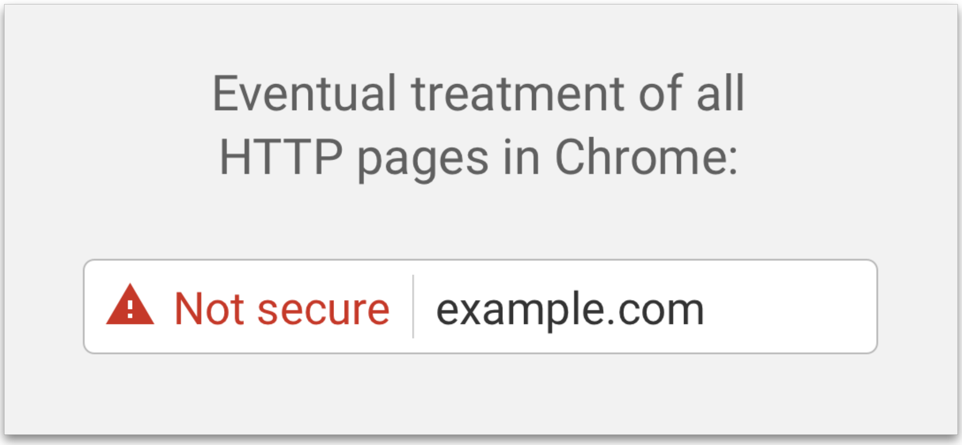
even when a certificate wasn't available. The motivation was to provide protection against passive surveillance (this would apply both to the teenager in the cafe and the NSA, in most cases) but without the more substantial guarantees from full HTTPS. In particular, that debate turned on whether Web site operators would consider the opportunistic encryption mode "good enough" and be disincentivized from providing additional security.[10]

Technical standardization proposals have also been used by parties opposed to the spreading of end-to-end encryption. A number of companies provide commercial services that depend upon inspecting and altering communications between Internet users: for example, anti-virus vendors or providers of exfiltration detection and prevention. These "middleboxes" want the capability to intercept these encrypted communications, decrypting them upon receipt, doing inspection for malicious attacks or the departure of sensitive data, and then re-encrypting them. While some end-to-end encryption proponents simply object to this model at all (given the potential for abuse of employees and customers, or alternative methods to achieve those security goals), some vendors have proposed standard ways for explicitly including a proxy as a party to the encryption, breaking end-to-end confidentiality, but maintaining some level of transparency or integrity. As implemented, these architectural means can allow for the continued operation of certain middlebox business models; they also serve a persuasive purpose in trying to promote alternatives that aren't fully end-to-end encrypted, or to provide a negotiating position that end-to-end encryption will be broken in various contexts.

Browser vendors can also use user interface design (which is typically explicitly not standardized across browsers) as an incentive for site operators to adopt security measures. These changes are typically made gradually, but Chrome has also signaled that it will eventually treat Web pages loaded over HTTP as explicitly "Not secure."

That red warning triangle might indicate to users something about the security situation that has long been normal, that there was no technical guarantee. More important for the purposes of this transition, it also provides a visual discriminator that might encourage users who are comparing sites to be cautious or wary of sites that are HTTP only. In that way, the code delivered to the many users of Google Chrome (on in this case, the blog post announcing some future changes in code) can affect market incentives.

[10] Would users be given some UI feedback that the channel was encrypted? If they were, it could more feasibly provide that disincentive for site operators. As some argued, even if users never realized that there was some additional level of security, they could still benefit from it.



Eventual treatment of all HTTP pages in Chrome:

⚠ Not secure | example.com

Figure 1: Screenshot of eventual treatment of HTTP in Google Chrome (Schechter 2016).

Norms

> The IETF community's technical assessment is that PM [pervasive monitoring] is an attack on the privacy of Internet users and organisations. The IETF community has expressed strong agreement that PM is an attack that needs to be mitigated where possible, via the design of protocols that make PM significantly more expensive or infeasible. (Farrell and Tschofenig 2014)

These are strong, blunt statements from a technical standard-setting organization. While direct about the values implicated (privacy), the framing is also limited in discussing "technical assessment" and denoting "attack" as a technical term rather than a judgment of malice. Similarly, while this is a community call for addressing surveillance in the design of standards, it is not as strict about specific conclusions as it might be. (There have been discussions of a "no new plaintext" document, but no such strict policy statements have been published.) Recognizing a consensus and describing "strong agreement" among a group is one way to document and encourage a change in norm.

Laws State actors notably have access to another mode of regulation of action; they can pass laws, rules and regulations and use law enforcement and penalties to encourage compliance. US intelligence agencies have repeatedly called for laws that would more explicitly restrict use of encryption so that wiretapping of Internet communications for law enforcement investigations would be easier. Legislative proposals from the FBI in May of 2013, for example, would have added financial penalties of $25,000 a day for Internet companies that did not successfully provide wiretap capabilities (Savage 2013).

This is another phase of the "Crypto Wars,"[11] a popular term used to describe debates between law enforcement and Internet companies and civil liberties advocates over the accessibility of encryption to the public. While these are debates over potential legislation, we might also interpret the very public statements of government officials as attempts to influence the norms of design of Internet communications technology.

Using handoffs

What do we gain from the handoffs model of analysis for the shift to encrypting Web traffic?

Identifying the handoff in values provides some protection against the naive assumption that a value simply didn't exist or wasn't provided prior to its technical implementation. Confidentiality of communications existed prior to TLS or to HTTPS-everywhere, it was just an unsteady assurance, provided by a mix of legal, social and market incentives. Identifying a trigger and a new paradigm provides a richer explanation of why this massive re-engineering of a system took place rather than a purely technical one: that a value wasn't present before, and now suddenly was.

In some ways the handoff here is straightforward, and may be a model for security features in many cases: discretion and responsibility is being removed from the end user (or some uncertain assumptions about other participants) and enforced cryptographically. To the end user, this might just appear like simple progress: if only more responsibilities for security vulnerabilities could be taken out of our hands (less constant vigilance about lock icons required, say) and instead guaranteed technically.

But how the handoff is actually accomplished is more complex: it relies on the coordination of many different actors – Web server operators around the world, notably, among others – and a combination of norms, market forces and architectural changes developed the path to the new paradigm. We can look at

[11] Or perhaps, as new proposals are about the re-designing of technology altogether, the "Design Wars" (Deirdre K. Mulligan and Doty 2016).

handoffs as shifting responsibility for a value, but also a triggering event and actions not just within each static paradigm but the modalities that move the socio-technical system between them.

When we apply the same model to Do Not Track,[12] we'll see a different handoff (not just human vigilance to security guarantee) but also a different set of actions within and between paradigms.

References

Bahajji, Zineb Ait, and Gary Illyes. 2014. "HTTPS as a Ranking Signal." *Google Search Central Blog*. https://developers.google.com/search/blog/2014/08/https-as-ranking-signal (November 29, 2020).

Chung, C., A. Kasyanov, J. Livingood, N. Mody, and B. Van. 2011. *Comcast's Web Notification System Design*. RFC Editor. https://www.rfc-editor.org/info/rfc6108 (November 29, 2020).

Clark, D D, and D R Wilson. 1987. "A Comparison of Commercial and Military Computer Security Policies." *IEEE Symposium on Security and Privacy* 0: 184–94. http://www.computer.org/portal/web/csdl/doi/10.1109/SP.1987.10001.

Farrell, S, and H Tschofenig. 2014. *Pervasive Monitoring Is an Attack*. RFC Editor. RFC. http://tools.ietf.org/html/rfc7258.

Lessig, Lawrence. 1999. *Code and Other Laws of Cyberspace*. Basic Books. http://books.google.com/books?id=0l1qLyT88XEC.

Mulligan, Deirdre K., and Nick Doty. 2016. "Design Wars: The FBI, Apple and Hundreds of Millions of Phones." *Citizen Technologist: The CTSP Blog*. https://ctsp.berkeley.edu/design-wars-fbi-apple/ (December 17, 2020).

Mulligan, Deirdre K, and Helen Nissenbaum. 2020. "Handoffs." *In progress*.

Okumura, Kaori, Yoshiaki Shiraishi, and Akira Iwata. 2013. "Survey on Sense of Security for Registering Privacy Information to Return Refugee Supporting System." In *Symposium on Usable Privacy and Security (SOUPS)*, https://cups.cs.cmu.edu/soups/2013/trustbusters2013/Sense_of_Security_Refugee_Supporting_System_Okumura.pdf.

Savage, Charlie. 2013. "U.S. Weighs Wide Overhaul of Wiretap Laws." *The New York Times*. https://www.nytimes.com/2013/05/08/us/politics/obama-may-back-fbi-plan-to-wiretap-web-users.html (December 17, 2020).

[12] See Do Not Track, a "handoff".

Schechter, Emily. 2016. *Moving Towards a More Secure Web*. https://security.googleblog.com/2016/09/moving-towards-more-secure-web.html (October 29, 2017).

Sunshine, Joshua, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. 2009. "Crying Wolf: An Empirical Study of SSL Warning Effectiveness." In *USENIX Security Symposium*, 399–416.

Thomson, Martin. 2014. *A Statement*. Internet Engineering Task Force. Internet-Draft. https://datatracker.ietf.org/doc/html/draft-thomson-perpass-statement-01.