Hacking the Anti-Hacking Statute: Using the Computer Fraud and Abuse Act to Secure Public Data Exclusivity

Nicholas A Wolfe
INTRODUCTION

Work smarter, not harder. Perhaps no other saying better captures the era of hyper-productivity and automation in which we live. Titles such as ‘Top Ten Hacks to Avoid Paywalls,’ ‘Five ways You’re Wasting Your time,’ and ‘One Weird Trick’ fly across our computer screens on a commoditized basis. These tips and tricks help us automate our lives and get more done, faster. Better living through automation. However, as these shortcut solutions get better and automation advances, a question arises. When does working smarter cross the line into cheating?

The Computer Fraud and Abuse Act was designed to draw this line, but in terms of hacking and computer crimes. Drafted in 1986 and amended on a frequency similar to iOS updates, the Act is intended to deter the exploitation of computer system vulnerabilities, but in so doing has exposed its own vulnerabilities in the face of a dynamic technological landscape. Vulnerabilities that are not better demonstrated than in the 3Taps decision. Additionally, this decision also highlights the manner in which companies are hacking the Act for use as a para-copyright tool to secure exclusivity to publicly-accessible data. This vulnerability arises in two respects: (i) the literal application of access controls and (ii) the cursory review of loss declarations pursuant to subsection 1030(e)(11). These vulnerabilities are clear when applied to public data. This Article therefore proposes that courts should interpret the Act in the context of data types and with reference to computer security technology.

I. WEB SCRAPING

* Legal Counsel at BitTitan, Inc.
2 Craigslist, Inc. v. 3TAPS INC., 942 F. Supp. 2d 962 (N.D. Cal. 2013)
Web scraping enables the copying and transposing of massive data sets into machine-readable and analyzable formats. Like web crawling (the familiar methodology employed by search engines), web scraping involves programmatic browsing of web pages and targeted data collection. Moreover, both methods collect data through end-user accessible network ports. Unlike web crawling, however, web scraping employs additional technologies to mimic human browsing and delve deeper each website. Innocuous web scraping technologies like the browser add-on that enables users to find free Pacer resources represent a bright spot in web scraping and its potential to advance the progress of science and the useful arts. Exploitative forms such as device or user impersonation designed to obtain private (often confidential) data, on the other hand, are less than inspiring, but nevertheless can be isolated to the context of private, confidential data.

II. THE COMPUTER FRAUD AND ABUSE ACT

Perhaps the most illustrating features of the CFAA are its amendments. In light of these many amendments and the difficulty inherent in applying a criminal statute to dynamic technology, there are nine ways to violate the CFAA. Fortunately for the reader, this article focuses on only one, which represents the civil-claim piece of the CFAA as well as the “lowest common

---


5 Id.


8 Making the writing of law review articles, for example, faster and less expensive.


denominator” in terms of prosecution. Subsection 1030(a)(2)(C) provides that anyone who (1) intentionally access without or in excess of authorization (2) information (3) that causes a plaintiff at least $5,000 loss in a 1-year period is engaged in hacking under the CFAA.

A. Unauthorized Access

The first element of subsection 1030(a)(2)(C), unauthorized access, is not clearly defined in the statute. Instead, the CFAA focuses on authorization in terms of permission and the scope thereof. Thus, not surprisingly some courts tend to interpret this element in terms of trespass doctrines. Accordingly, this interpretation leads courts to apply policy-based safeguards, such as terms of use, computer use restrictions, and even an employee’s general duty of loyalty, to trigger liability under the CFAA. Notably, the Seventh, Fifth, and Eleventh Circuits still follow this broad interpretation. In Brekka, the Court held that an employee’s duty of loyalty does not apply to whether such employee (or former employee) accessed information without authorization under the CFAA. The Court reasoned that this interpretation inappropriately shifts the element to the mental state of the violator, as opposed to the nature of access. In Nosal, the Court further narrowed the scope of unauthorized access to better align with legislative intent by holding that unauthorized access (a) applies only to the circumvention of technological safeguards and (b) does not apply in the context of use controls. Concerning the circumvention of technological safeguards, the Court reasoned that narrowing application of the CFAA to circumvention of technological barriers better aligns the statute with its anti-hacking intent and cited its decision in Brekka for support. Moreover, concerning use controls, the Court used the online dating scene to illustrate the untenable impact of importing use controls CFAA liability. Somewhat playfully, the Court referenced the perpetually 29-year old online profiles in violation of applicable terms of use and the widespread possibility of punishing innocuous use-related activity as hacking. Perhaps more impactful than the relief provided to online singles everywhere, the court noted that the CFAA was not intended to apply as a misappropriation tool.

13 Id.
In 3Taps, the defendant used a proxy server in circumvention of an IP block initiated on otherwise publicly-accessible data. In so doing, the 3Taps decision construed the Nosal decision to encompass perfunctory safeguards.\textsuperscript{14} Grouping perfunctory, or mere token, safeguards with functional safeguards exposes a "hack" in the framework of the statute and Nosal. Nosal stands for the proposition that use controls are beyond the scope of the CFAA.\textsuperscript{15} Further, in order to be defined as an access control it follows that the control would need to provide some access-restricting function.\textsuperscript{16} Otherwise, a danger arises that a plaintiff need only initiate patchwork, token access controls to disguise a use control on public data.\textsuperscript{17} In support, relevant authority reveals that the lower sophistication of the safeguard involved, the higher tendency there is for using the CFAA as a misappropriation tool. For example, in holding that the defendant’s violations of terms of use cannot be defined as hacking under the CFAA, the court in Cvent reasoned that the plaintiff’s data is “not protected in any meaningful fashion by its Terms of Use or otherwise.”\textsuperscript{18}

Nevertheless, the 3Taps court took a page from the Seventh Circuit’s reasoning in Citrin and found solace in the fact that the defendant had clear notice that its access was unwanted.\textsuperscript{19} Similar to the homeowner who uses an ADT sticker to safeguard herself from home invasion, the court shifted its focus form functional web security to notice; a shift that does not square with the statute’s access-centered, anti-hacking intent. Notably, had the defendant decided to read Craigslist advertisements in Spanish, it would have been engaged in hacking under this logic.\textsuperscript{20}

\textbf{B. Information}

Perhaps the most critical step in the 3Taps Court’s reasoning was “[a]ssuming that the CFAA encompasses information generally available to the public.”\textsuperscript{21} Taken by itself, this step is a logical step necessary to the conclusion drawn by the Court. However, in the context of the CFAA, this

\textsuperscript{14} See 3TAPS, 942 F. Supp. 2d at 970.
\textsuperscript{15} Nosal, 676 F. 3d at 863-864.
\textsuperscript{17} See Nosal, 676 F. 3d at 857.
\textsuperscript{18} Cvent, 739 F. Supp. 2d at 933 (emphasis added)
\textsuperscript{19} 3TAPS, 942 F. Supp. 2d at 970.
\textsuperscript{20} The use of Google Translate software masks the originating IP address similar to using a proxy server.
\textsuperscript{21} 3TAPS, 942 F. Supp. 2d at 969.
assumption has widespread import. For example, the CFAA does not define the meaning of “information.” Congressional history indicates that the CFAA was designed to protect the privacy of data. In amending the CFAA to include subsection 1030(a)(2)(C), drafters stated that “[t]he proposed subsection 1030(a)(2)(C) is intended to protect against the interstate or foreign theft of information by computer.” Further drafters stated the entire “premise of this subsection is privacy protection” and that unauthorized access under this subsection “includes mere observation of the data.” These three elements - theft, privacy, and mere observation – cut against the assumption that subsection 1030(a)(2)(C) applies to public data.

The dichotomy between CFAA treatment of public and private data becomes even more pronounced in judicial application. In *Facebook II*, the defendant Power Ventures developed scraping software whereby users input their Facebook login credentials to “invite their friends” to join the Power Ventures’ service offering. The software accessed the user’s Facebook account and email contact list, where it then sent invite emails to the private contact list without recipient permission. Importantly, scraping of private data can generally be distinguished from scraping of public data in that the former involves circumvention of individualized login or some other type of obfuscation applied against all users. For instance, in *EF Cultural Travel*, the First Circuit held that the use of a web scraper in combination with information not readily accessible to users amounts to unauthorized access under the CFAA. Here, the defendant used the services of a current employee to translate tour codes for use in the web scraping software. The software would then access the website and submit tour codes and harvest related pricing. In isolation, the tour codes were meaningless. However, when combined with the knowledge of how EF Cultural Travel’s website used the tour codes in connection with querying a backend database for current pricing, use of the tour codes

---

24 *Id.*, at 6
25 *Id.*
26 *Facebook, Inc. v. Power Ventures, Inc.*, 844 F. Supp. 2d 1025, 1028 (N.D. Cal. 2012) (Facebook II)
27 *Id.*
28 See *EF Cultural Travel BV v. Explorica, Inc.*, 274 F. 3d 577, 579 (1st Cir. 2001)
29 *Id.*
30 *Id.*
became more exploitative in nature.\textsuperscript{31}
Notably, the 3Taps court did not cite any precedent applying the CFAA to public data.\textsuperscript{32} There are very few judicial applications to public data in the context of the CFAA. In Register.com v. Verio, the defendant developed a scraping software to harvest publicly-accessible WHOIS domain registrant email addresses.\textsuperscript{33} Verio emailed registrants for the purposes of offering a competitive website development service.\textsuperscript{34} Register was an ICANN registrar and operated the public database pursuant to an agreement with ICANN and its own terms of use.\textsuperscript{35} The court focused on Verio’s violation of Register’s terms of use in the context of contract law, and declined to weigh in on the applicability of CFAA.\textsuperscript{36} However, the dissenting opinion makes clear that establishing damages in this context is improbable.\textsuperscript{37} Moreover, terms of use that purport to restrict the use of public data, but which do not require a private, individualized step to limit access are unenforceable in many jurisdictions and raise concerns equivalent to those encountered in Feist regarding protection of factual, public data.\textsuperscript{38} For instance, in Ticketmaster, a California federal district court examined the potential misappropriation claims of public data in the context of preemption under the Copyright Act.\textsuperscript{39} Although not applying the CFAA, the Court emphasized the danger for protecting factual, public data under paracopyright claims in circumvention of the Copyright Act and the longstanding principle in Feist.\textsuperscript{40}

If anything, the dearth of CFAA application to protect publicly-accessible data suggest its inappropriateness under an anti-hacking statute. Notably, Senator Leahy recently proposed an amendment to limit the

\textsuperscript{31} Id., at 582-583
\textsuperscript{33} Register.com, Inc. v. Verio, Inc., 356 F. 3d 393, 395-396 (2nd Cir. 2004).
\textsuperscript{34} Id., at 395.
\textsuperscript{35} Id.
\textsuperscript{36} Id., at 401-402.
\textsuperscript{37} Id., at 440 (Parker, J. dissenting) (“To maintain a cause of action under the CFAA against Verio, Register.com must demonstrate the Verio violated the CFAA in a manner that has caused Register.com damages or losses of at least $5,000. There is nothing in the record to suggest that this has occurred.”)
\textsuperscript{39} Id.
\textsuperscript{40} Feist, 499 U.S. at 363
information protected under the CFAA to seven categories of private information, such as passwords or personally-identifiable information.41

C. Establishing Loss

In order to maintain a civil action against a defendant under the CFAA, a plaintiff must demonstrate a loss to 1 or more persons during any 1-year period aggregating at least $5,000 in value.42 A “loss” is defined as any reasonable cost to any victim, including the cost of responding to an offense, conducting a damage assessment, and restoring the data, program, system, or information to its condition prior to the offense, and any revenue lost, cost incurred, or other consequential damages incurred because of interruption of service.43 Prior to this definition being added to the statute in the USA PATRIOT Act of 2001, courts had difficulty applying in the context of business entities.44 Eventually, the Ninth Circuit solved the problem by defining “loss” in terms of system-related harm, such as service interruption or requiring users to change passwords.45 For example, courts have held that the mere copying of proprietary data is insufficient to establish a loss under the CFAA.46 Instead, a plaintiff must show that there was an impairment to its computer system or data as a result of the intrusion. Moreover, a plaintiff must connect the loss to the actual intrusion.47

Congress also intended that the loss be reasonably or necessarily incurred by the plaintiff.48 Jury instructions often track this requirement closely.49 There are three scenarios that are applicable to web scraping cases, each of which demonstrate the requirement that a loss must be reasonably incurred. The first is with respect to spamming. For example, the court in Facebook II focused on the potential violations of the CAN-SPAM Act to substantiate the plaintiff’s argument that it had properly established the loss requirement

43 Id., at 1030(e)(11).
45 United States v. Middleton, 231 F. 3d 1207, 1213 (9th Cir. 2000).
49 Middleton, 231 F. 3d at 1213 (“In determining the amount of losses, you may consider what measures were reasonably necessary to restore the data, program, system, or information that you find was damaged or what measures were reasonably necessary to resecure the data, program, system, or information from further damage.”).
under the subsection 1030(e)(11). At issue in *Facebook II* was the defendant’s software that requested user credentials, and then impersonated a user and emailed their Facebook contact lists with unauthorized communications. The second scenario is where a password portal has been circumvented. In *Successfactor*, the Court sided with the plaintiff in its argument that when a password protected environment is bypassed, certain costly security measures need to be undertaken to ensure network and application integrity. Moreover, in *Vanderhye v. iparadigms*, the Fourth Circuit held that investigation into the possibility of a technical glitch in the system qualified as a loss under the CFAA, because such glitch impacted the way in which the plaintiff secured private data. Notably, each of these cases involved a password-protected portal. There is also a third scenario which demonstrates that a loss must be reasonably incurred, but does so in the negative. In the famous case *Bidder’s Edge*, the plaintiff claimed harm to its system based on the potential for unchecked aggregation of access and irreparable harm caused thereby. In overruling this decision, the Court in *Intel v. Hamidi* held that the potential for harm is not sufficient proof of harm in terms of trespass to chattels. The Fifth Circuit follows suit, stating that web scraping of a publicly-available website, without more, is insufficient to fulfill the harm requirement (in the context of trespass to chattels).

In *3Taps*, Craigslist loss declaration cites an internal investigation into the harm caused by the Defendant’s intrusion, well in excess of the $5,000 threshold. The Court’s primary interest centered on whether Defendant had access the information without authorization, without nothing more than a cursory glance at the nature of the loss involved. Notably, California Penal Code section 502 does not require a similar showing of loss.

What is interesting about the *3Taps* decision is that the Court dealt with the CFAA application to public data with nothing more than literal application of the statute; whereas, the decision discussed previously devote considerable more attention to the nature of the data and the loss incurred in maintaining its security. Moreover, outside of *3Taps*, the case history on whether the use of a proxy server is harmful to a computer is nonexistent.

---

50 See e.g., *Facebook II*, 844 F. Supp. 2d at 1028; *America Online, Inc. v. LCGM, Inc.*, 46 F. Supp. 2d 444 (E.D. Va. 1998).
52 *AV ex rel. Vanderhye v. iparadigms, LLC*, 562 F. 3d 630, 645 (4th Cir. 2009).
55 *White Buffalo Ventures, LLC v. University of Texas*, 420 F. 3d 366 (5th Cir. 2005).
proxy server does not change the nature of access, but instead changes the
identity of the person accessing, which is a critical distinction in the context
of public data. This means that had the Defendant translated a Craigslist
advertisement in Spanish, it would have caused the same harm to Craigslist
computers. Releasing data to the public has certain inherent risks, risks that
are not affected by the identity of the person accessing. That Craigslist
needed to investigate the use of a proxy server seems disingenuous at best.
At worst, such investigations represents the perfunctory circumvention of a
meaningful statutory requirement. Notably, Craigslist did not require
passwords or the CAPTCHA requirement used by many U.S. government
websites to prevent robotic access.57

PART IV. CHALLENGES AND PROPOSAL

The arguments advanced in this article are not intended to require a
plaintiff under the CFAA to have effective safeguards or access controls.
Indeed, the effectiveness of web security is in the eye of the hacker, and not
all hackers are created equal. Nor does the author intend that the CFAA
should not be applied to public data which are defaced, destroyed, or
damaged by hacking attacks. Rather, the intent of this article is to question
the judicial application of meaningful statutory requirements in a cursory
manner.

Recent legislative proposals aim to solve this problem. Aaron’s Law
represents a step in the right direction, but also contains a flaw common in
the many amendments to the CFAA.58 Applying a hacking statute to cover
specific scenarios is necessary backward-looking in the fact of dynamic
technology. In other words, past trends do not indicate future results.

Instead, Senator Leahy’s proposal to identify seven categories of
protected information represents the best solution to the vulnerabilities in the
CFAA. Not only does this proposal avoid the CFAA being used as a
misappropriation tool, but it prevents the CFAA from being applied to punish
innocuous activity, which in some cases represents advancement of the
sciences and useful arts.

During the interim, courts should be mindful of the potential wielding of
the CFAA as a sword through cursory, literal compliance with the statutory
elements. Circumvention of access controls should be interpreted in terms
of web security as opposed to mere unwanted access. Intrusion-based losses
should be reasonably incurred and read in light of risk already inherent in the

57 See United States Patent and Trademark Office, Public PAIR,
58 H.R.2454, 113th Cong. (“Aaron’s Law Act of 2013”)
presentation of the plaintiff’s data. Finally, account should be made of the data type involved. Applications of the CFAA to public data generally warrant more scrutiny in order to avoid untenable results.

For example, in the case of a news website that uses a paywall to restrict access. If a user accesses the website in excess of her pre-defined article limit, a notice to pay is presented in lieu of access to news. Without more scrutiny, a user on a Google Chrome browser, who presses “Ctrl + U” is likely now a hacker under the CFAA.²⁹

***

²⁹ Note that Ctrl + U presents HTML Code format of the web page, and allows a user to circumvent and HTML-based paywall.