Harvard University | Data Privacy Lab

Latanya Sweeney, PhD
Professor of Government and Technology in Residence
Director, Data Privacy Lab
Institute for Quantitative Social Science

1737 Cambridge Street, K310
Cambridge, MA 02138
(617) 496-3629 latanya@fas.harvard.edu
latanyasweeney.org | dataprivacylab.org

June 24, 2019

The Honorable Mikie Sherrill, Chairwoman
The Honorable Haley M. Stevens, Chairwoman
U.S. House Subcommittee on Investigations & Oversight
U.S. House Subcommittee on Research & Technology
2318 Rayburn House Office Building
Washington, DC 20515

Re: Hearing on Election Security: Voting Technology Vulnerabilities

Dear Chairwoman Sherrill and Chairwoman Stevens:

I write to you in advance of the hearing on "Election Security: Voting Technology Vulnerabilities." I appreciate your interest in securing websites that maintain voter information. I was the lead author on a scientific paper that surveyed vulnerabilities in voter information websites in 2016[1]. My co-authors, Ji Su Yoo and Jinyan Zang, work with me on the Technology Science Initiative, in the Institute for Quantitative Social Science at Harvard University. We welcome your leadership on this critical issue and look forward to working with you and your staff.

In 2016, we conducted a series of scientific investigations into ways an attacker could use technology in an attempt to adversely impact elections. We found misinformation about polling place locations, which by November were corrected.

We also found websites for 35 states and DC in 2016 that were vulnerable to voter identity theft attacks: an imposter could submit changes to voter registration information. An imposter needed a combination of voter's name, date of birth, gender, address, Social Security Number, or Driver's License Number.

Relevant data could be acquired from government, data brokers, or darknet markets. Total cost of an automated attack against 1percent of all vulnerable voter registrations nationwide ranged from $10,081 to $24,926 depending on the data source used. States cost less, e.g., $1 for Alaska and $1,020 for Illinois.

A voter identity theft attack could disrupt an election by imposters submitting address changes, deleting voter registrations, or requesting absentee ballots.

[1] Sweeney L, Yoo J, Zang J. Voter Identity Theft: Submitting Changes to Voter Registrations Online to Disrupt Elections. *Technology Science*. 2017090601. September 06, 2017. Version 2. https://techscience.org/a/2017090601

Voter Identity Theft

Could an attacker impact U.S. elections by merely changing voter registrations online? This reportedly happened during the 2016 Republican primary election in Riverside County, California. What about elsewhere? We surveyed official voter record websites for the 50 states and the District of Columbia and assessed the means and costs for an attacker to change voter addresses. Relatedly, an attacker could also change party affiliations, delete voter registrations, or request absentee ballots online. A voter whose address was changed without her knowledge, for example, in most states would have a polling place different than expected. On Election Day, when she appeared at her presumed polling place, she would have been unable to cast a regular vote because her name was not on the precinct's register. She may have been turned away or given a provisional ballot, and in many cases, a provisional ballot would not count. Perpetrated at scale, changing voter addresses, deleting voter registrations, or requesting absentee ballots could disenfranchise a significant percentage of voters, and if carefully distributed, such an attack might go unnoticed even if the impact was significant. So, how practical is it to submit false changes to voter registrations online?

In summary, we found that in 2016, the District of Columbia and 35 of the 50 states had websites that allowed voters to submit registration changes. These websites determined whether a visitor was an actual voter by requesting commonly available personal information. Some websites gave multiple ways for a voter to self-identify. Of these, {name, date of birth, address} was required in 15, {name, date of birth, driver's license number} was required in 27, and {name, date of birth, last 4 SSN} was required in 3. We found that an attacker could acquire the voter names, demographic information and government-issued numbers needed to impersonate voters on all 36 websites from government offices, data brokers, the deep web, or darknet markets.

Overall, the total cost of an attack in 2016 varied based on the number of voters to impersonate, data sources used, whether the websites had CAPTCHAs, and specific states of interest. We found that the practical costs of changing 1 percent of the voters on all 36 websites could range from $10,081 to $24,926 depending on whether the attacker used data from government, data broker, darknet or other sources. Costs for an attack on a specific geographical area or state were much less, such as $1 for Alaska or $1,020 for Illinois. Back office processes and election practices, which varied among states, could have possibly limited attack success rates.

Fundamental Cybersecurity Vulnerabilities

Usually "cybersecurity" focuses on ways an attacker can break into a system or steal the credentials of those administrators and officials who use the internals of the system. Once inside, the attacker has open access to the files and systems. For this reason, perimeter security that surrounds the stored information is critical. These can be addressed through traditional computer security best practices, including but not limited to those proposed by the Help America Vote Act (HAVA), the Voluntary Voting Systems Guidelines (VVSG) by the National Institute of Standards and Technology (NIST) and the Election Assistance Commission (EAC).

However, I want to point out that many government websites have unique security concerns that go beyond the ability to secure the perimeter. Additional vulnerabilities exist because the intended users of many government systems are members of the public who identify themselves to the systems using personal information that is also widely available. For example, the State of Delaware had a website for voter's to change their voter registration information online. A voter identifies himself to the system using {name, date of birth, 5-digit ZIP code}. The voter knows this information, but unfortunately, we showed that this same information was readily available from voter lists, data brokers and on the dark web. An attacker could impersonate a voter at scale on these websites to impact elections. Different state websites used different combinations of personal demographics and government issued identifiers, including Social Security numbers and driver's licenses. But all the combinations of information requested were available to an attacker. Even with perfect perimeter security afforded by traditional cybersecurity, an attacker could still commit "voter identity theft" and change voter records at scale through automated means inexpensively.

Assistance Congress Could Provide to Assist States and Counties in Securing Websites

Our findings identify the nature of the problem, but they also suggest best practices to limit or thwart voter identity theft.

In our paper, we computed the costs of changing one percent of the voter records at each website. Costs included the acquisition of the specific pieces of information needed to impersonate voters at the state website and the costs of using virtual machines to automatically change different records slowly over time to avoid human detection. The costs varied significantly among the states: Alaska was $1, Delaware was $7, and Ohio $330, as examples. The most expensive state was Texas at $3,059. The key characteristic that the Texas website had that made it more difficult to impersonate its voters was a serial code that appeared on the face of a driver's license that could not be computed from the demographics itself. Texas voters had to enter this code, but this code was not available from data brokers who provided driver license numbers. Impersonating Texas voters online required images of actual Texas driver's licenses, which we did find on the darknet. Clearly, using this number helps thwart identity theft.

One of the reasons automated attacks were inexpensive was because few websites had those annoying pop-up boxes that attempt to stop automation. CAPTCHAs as they are termed, request selecting a subset of images, entering text from an image, or performing some other task that should be easy for a real human to perform but difficult for an automated script to achieve. CAPTCHAs help defeat voter identity theft by limiting the speed of how many voters could be impersonated in a time period.

Eleven (31 percent) of the 36 websites we found in 2016 had a CAPTCHA service. But automated programs could respond to the kinds of CAPTCHAs found on all the state websites that had CAPTCHAs, thereby rendering them a nominal deterrent. Improvements have been made in recent CAPTCHAs.

My colleagues and I urge the Subcommittees to explore ways to help state and county websites use special codes that may appear on driver licenses and to use the latest versions of CAPTCHAs on websites that allow voters to change voter information.

My colleagues and I also urge the Subcommittees to provide research funds to develop anomaly detection algorithms on voter data so that unusual activity can be identified, and alerts sent to officials for human inspection. These alerts can identify an assortment of problems, even violations that come from penetration of the perimeter security. (In the interest of full disclosure, my colleagues and I have begun such an effort.)

I also want to make a distinction that the websites having the vulnerabilities we describe are websites that allow voters to change their voter information. Sometimes, these were voter registration websites, but other times, they were motor vehicle websites that did not even allow new voters to register to vote but did allow voters to change existing registrations.

My colleagues and I are busily re-surveying the state websites now to provide updated information. When these results are finalized, we will forward them to you.

I ask that this letter be entered in the hearing record. My colleagues and I look forward to working with the Subcommittees on these issues of vital importance to the American public.

Yours truly,

Latanya Sweeney, PhD.