Security

Computer forensics defuses FBI's Clinton email 'bombshell' – the math doesn't add up

95 per cent of the 650,000 messages not relevant

4 Nov 2016 at 16:56, Duncan Campbell 44

Analysis Since igniting a political firestorm and triggering major changes in US presidential voting intentions by revealing some emails passing through Hillary Clinton's private email server had been found in an unrelated criminal investigation, the FBI has gone to ground.

The US criminal investigation bureau has repeatedly refused to answer basic media questions about simple and long-established computer forensic procedures.

But the math, based on detailed information previously released by the FBI, points to the conclusion that the agency will have known by Monday morning exactly how many emails found in a laptop computer seized a month ago from disgraced former New York Congressman Anthony Weiner had come from, gone to, or been copied on from the Clinton server, and how many, if any, could contain possibly classified information not already checked.

The agency appears to have pushed a completely misleading number out to US media outlets, suggesting that 650,000 emails had to be checked.

Comey told Congress: "The FBI cannot yet assess whether or not this material may be significant. I cannot predict how long it will take to complete this additional work."

But the FBI did not point out that of the 650,000 emails mentioned to the US media, 95 per cent could not possibly be relevant.

Comey's letter to Congressional leaders, which started the whole debacle, explained that the agency could not officially look at or report on the emails without obtaining a specific new warrant. The letter implicitly acknowledged that the agency already had copies of all the mails on its computer systems

More like this

Fbi Donald Trump

Most read


DRAMA ON MARS: Curiosity bot fires laser at alien metal object


Amazon's very own Linux now available for download


Leaks password, check. Leaks Wi-Fi password, check. Can be spoofed, check. Ding! We have an Internet of S**t winner


A British phone you're not embarrassed to carry? You heard that right


Why Apple's adaptive Touch Bar will flop

Spotlight

(which would normally automatically have been indexed by forensic software), bringing the Clinton connection to light.

To find out how many emails on the laptop were relevant would have taken "seconds", according to e-discovery software industry experts. To then find out how many of those – if any – the FBI had not seen in its previous investigation would, at most, have taken "minutes." Standard methods are to take and match cryptographic hashes of email files (which proves the email files identical, if the hashes match), or to match metadata and then textual content.

The FBI's previous, year-long investigation into the private Clinton server finished in July, when director James B Comey reported that: "We cannot find a case that would support bringing criminal charges."

As only 110 of 30,490 official emails previously examined by the FBI were found to contain classified government information, the number of previously unseen mails that had strayed onto Weiner's laptop is likely to range from zero to a few tens.

How the mess began

The laptop at the heart of the election controversy was seized on October 3 from former Congressman Weiner after a then-15-year-old girl from North Carolina had complained of sexting. The alleged victim, now 16, has now complained vociferously that Comey had irresponsibly forced her identity into the open, exposed her to continual and continuing media harassment, and caused the abuse to continue.

"You have assisted him in further victimizing me on every news outlet. I can only assume that you saw an opportunity for political propaganda," she said.

Standard forensic procedures for e-discovery in civil and criminal investigations is to make a certifiable digital copy of all media immediately after getting access, and immediately to analyse and index the contents, including buried metadata and email attachments.

The software utilised in these investigations is used to handling and sifting big data, scaling up to tens of millions of files. The global e-discovery market in software systems and services is now worth an estimated $1bn, with many companies offering sophisticated email analysis add-on systems to spot, map, network and visualise chaining, duplicates, and to provide searchable indexes.

The FBI have long been leaders in this business. As revealed by Edward Snowden, the FBI has been operating the PRISM and other systems for over ten years from its Digital Intercept Technology Unit (DITU) at its sprawling Quantico, Virginia base. The unit annually "ingests" and analyses billions of emails intercepted from US optical fibre cables or passed on by telecommunications operators. The critical part of the system's front end, obviously, is to spot email addresses associated with intelligence targets.

But when it came to the debate, the agency's computer teams had apparently regressed to the digital stone age. *The New York Times* reported: "The FBI needed custom software to allow them to read Mr Weiner's emails without viewing hers. But building that program took two weeks."

Industry experts used to massive email searches in large civil cases have been scathing about the idea that the FBI's job is difficult with modern tools. Linda Sharp of ZL Technologies said: "In the scheme of e-discovery, 60,000 documents is nothing. We're used to seeing documents in the tens of millions of documents, terabytes of data."

Even if you read every email, "we're not talking about a lot. 60,000 is nothing."

Journalists have also become users of high-end e-discovery software to handle document dumps in recent high profile reports, such as the Panama Papers and Offshoreleaks investigations (Duncan worked as the data manager for the Offshoreleaks project of the International Consortium of Investigative Journalists). In the Offshoreleaks investigation in 2013, two million emails were analysed and catalogued, and made available to international journalism teams on a secure server. To find all emails from a domain takes seconds, once the gruntwork of indexing is complete – which had previously been done for Weiner's computer, to look for sexing evidence.

Standard WHOIS registry records show that the clintonemail.com domain was registered on 13 January 2009. She turned down the opportunity to use a standard state.gov address, and corresponded throughout her term of office as hdr22@clintonemail.com.

In 2009, Clinton appointed Huma Abedin as deputy chief of staff at the State Department. In 2010, Abedin married Weiner. They separated this past August. Abedin then became vice chairwoman of

Good luck securing 'things' when users assume 'stuff just works'

Democralypse Now? US election first battle in new age of cyberwarfare

You've been hacked. What are you liable for?

Securing Office 365? There's always more you can do

DDoS attacks: For the hell of it or targeted – how do you see them off?

Hillary Clinton's 2016 Presidential campaign. Apart from communicating with Clinton on her email, Abedin and another aide also had personal accounts on the Clinton server.

The implication of the FBI's October findings is that Abedin communicated with her husband from the clintonemail domain, or copied him some of her boss's email, or even that he lifted and copied them in a domestic setting.

Whichever happened, or all of them, finding those emails on Weiner's laptop will have been forensically trivial, as all will contain the unique string "clintonemail." Google it and you get it, in seconds.

Republicans have form for previously ~~exploiting~~ making fundamental forensic errors in reporting on email data in the Clinton investigation. In 2015, it was claimed that she had a second "secret" address on the server. In fact, it was a new address she used after being Secretary of State.

Phoney numbers

Asked by *The Reg* if they agreed that as their own investigation into Clinton reported that there were 62,320 emails handled on the clintonemail.com domain during her term in office as Secretary of State, and that they had already checked 30,490 of those handed over by her lawyers as being official, 90 per cent must be irrelevant – an FBI spokesman refused comment.

The Reg asked how long it had taken them to filter the emails to select only Clinton mails, and how many had actually been found. "No comment."

Do the math. The FBI have already seen nearly half of the emails handled by the server. The balance of emails deemed private by Clinton's lawyers is 32,740. Even if, implausibly, the entire contents of the Clinton server had been copied to Abedin, and then on to Weiner, it is obvious that 95 per cent of the Weiner emails could not be relevant. Commonly, two such troves contain many sets of multiple copies of the same emails, made automatically by backup and other processes.

Oregon Senator Ron Wyden, a longstanding critic of FBI and NSA electronic mass surveillance, told *The Reg* that the FBI's "continuing leadership failures" underscore the "need for independent oversight" on surveillance, and reflected a "pattern of poor judgment" by the FBI's director.

The US media have been full of hyperbole about how no effort has been spared by the FBI in its efforts to break the butterfly on their wheel. They would "spare no resources," are working "round the clock" on "16-hour shifts," developing "new software" for the taxing task.

In an internal FBI message reported by NBC, Comey is said to have told agents that it would "be misleading to the American people were we not to supplement the record. At the same time, however, given that we don't know the significance of this newly discovered collection of emails, I don't want to create a misleading impression", he added. Indeed. ®

Sponsored: 10 Reasons LinuxONE is the best choice for Linux workloads

Tips and corrections 50 Comments

Whitepapers

Empowering E-commerce against unpredictable demand
Exploring common challenges facing E-commerce platform and B2C website managers planning for the holiday shopping rush and bolstering services to meet customer demands.

Data and Analytics Maturity Model and Business Impact
How top performing enterprises use their IT investments to store, process, and use data to make more effective, real-time decisions.

Controlling the Uncontrollable
Protecting email and other corporate data on mobile devices—without bogging down workers—is one of today's biggest challenges faced by IT pros today.

Email is the critical communications tool; interruptions are costly

National Cyber Security Centre to shift UK to 'active' defence

You call it 'hacking.' I call it 'investigation'

NHS health apps project plan: Powered by your medical records

Sponsored links

Sign up to The Register to receive newsletters and alerts

At LOGICnow, we have invested significantly to develop email security services that are reliable, scalable and secure.

The Register

Biting the hand that feeds IT © 1998–2016

Independent news, views, opinions and reviews on the latest in the IT industry. Offices in London, San Francisco and Sydney.