An Introduction of Cybersecurity and Flash Memory Security Research

Bo Chen
Assistant Professor, Department of Computer Science
bchen@mtu.edu
https://cs.mtu.edu/~bchen
https://snp.cs.mtu.edu
About Me

Bo Chen
Assistant Professor, Computer Science

Associate Professor effective in August 2022

Areas of Expertise
- Mobile Device Security
- Cloud Computing Security
- Named Data Networking Security
- Big Data Security
- Blockchain

Links of Interest
- Faculty Website
- MTU Security and Privacy (SnP) lab

https://cs.mtu.edu/~bchen

Lab director: MTU Security and Privacy (SnP) Lab

Co-advisor: MTU CS cybersecurity reading group, MTU RedTeam
Faculty coach: MTU NCL cyber competition team (our team ranked 10th out of 3916 teams across US in Fall 2021)
About MTU Security and Privacy (SnP) Lab

The Secure and Privacy (SnP) lab at Michigan Technological University was established in early 2018. The mission of SnP lab is to promote research and education of cybersecurity. For research, we aim to tackle cutting-edge security and privacy problems, protecting safety and assets of people from malicious attacks. For education, we are enthusiastic about broadcasting cybersecurity knowledge among graduate and undergraduate students. We are also dedicated to promoting cybersecurity training among underrepresented groups and future cybersecurity professionals through various outreach efforts.

https://snp.cs.mtu.edu

Lab Director
Bo Chen (Assistant Professor@MTU CS)

PhD Students
Niusen Chen
Md Mezbahul Islam
Wen Xie

Master Students
Thomas Grifka
Sai Venkateswaran

Undergraduate Students
Dominika Bobik
Ethan Brinks
Josh Dafoe
Ryan Klemm
DeAndre Neal (MiCUP student)
About MTU Security and Privacy (SnP) Lab

• Projects are currently funded by national science foundation, national security agency, etc.
  • Protecting sensitive data in mobile devices, IoT devices
  • Protecting critical data/infrastructures outsourced to public clouds
  • Blockchain and information centric networking
  • Malware detection
  • Security and privacy in connected and autonomous vehicles
A Starting Video

- https://www.youtube.com/embed/ThBpRBpyxLI?start=0&end=50&version=3
All Starts from Malware and Hacks

• On November 2, 1988, a graduate student at Cornell University, Robert Morris, unleashed what became known as the Morris worm
  • Morris worm disrupted a large number of computers then on the Internet, guessed at the time to be 10% of all those connected

• Malware and Hacks are here and there today

Total malware
How to Combat Malware and Hacks?

• The answer is cybersecurity

• Ensure our systems and networks are well protected, such that any intruders can be detected, identified, and blocked
  • Make sure the software (code) we build is free of vulnerabilities
    • The attackers cannot exploit the vulnerabilities to intrude into our systems and networks

• Make sure our data are protected
  • Not disclosed to unauthorized parties
  • Not modified by unauthorized parties
  • Always available for use
  • Always recoverable
Outline

• Recent cybersecurity instances
• What is cybersecurity
• Why learning cybersecurity
• Mobile devices and flash memory
• Flash memory security research
Recent Cybersecurity Instances
Colonial Pipeline attack: Everything you need to know

Updated: DarkSide has claimed responsibility for the catastrophic ransomware outbreak.
Hacks in 2020

Exclusive: Iran-linked hackers recently targeted coronavirus drugmaker Gilead - sources

Exclusive: Elite hackers target WHO as coronavirus cyberattacks spike
<table>
<thead>
<tr>
<th>Entity</th>
<th>Year</th>
<th>Records</th>
<th>Organization type</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adobe Inc.</td>
<td>2019</td>
<td>7.5 million</td>
<td>tech</td>
<td>poor security</td>
</tr>
<tr>
<td>Amazon Japan G.K.</td>
<td>2019</td>
<td>unknown</td>
<td>web</td>
<td>accidentally published</td>
</tr>
<tr>
<td>2019 Bulgarian revenue agency hack</td>
<td>2019</td>
<td>over 5,000,000</td>
<td>government</td>
<td>hacked</td>
</tr>
<tr>
<td>Canva</td>
<td>2019</td>
<td>140,000,000</td>
<td>web</td>
<td>hacked</td>
</tr>
<tr>
<td>Capital One</td>
<td>2019</td>
<td>106,000,000</td>
<td>financial</td>
<td>unsecured S3 bucket</td>
</tr>
<tr>
<td>Desjardins</td>
<td>2019</td>
<td>2,900,000</td>
<td>financial</td>
<td>inside job</td>
</tr>
<tr>
<td>DoorDash</td>
<td>2019</td>
<td>4,900,000</td>
<td>web</td>
<td>hacked</td>
</tr>
<tr>
<td>Facebook</td>
<td>2019</td>
<td>540,000,000</td>
<td>social network</td>
<td>poor security</td>
</tr>
<tr>
<td>Facebook</td>
<td>2019</td>
<td>1,500,000</td>
<td>social network</td>
<td>accidentally uploaded</td>
</tr>
<tr>
<td>Facebook</td>
<td>2019</td>
<td>267,000,000</td>
<td>social network</td>
<td>poor security</td>
</tr>
<tr>
<td>First American Corporation</td>
<td>2019</td>
<td>885,000,000</td>
<td>financial service company</td>
<td>poor security</td>
</tr>
<tr>
<td>Health Sciences Authority (Singapore)</td>
<td>2019</td>
<td>808,000</td>
<td>healthcare</td>
<td>poor security</td>
</tr>
<tr>
<td>Justdial</td>
<td>2019</td>
<td>100,000,000</td>
<td>local search</td>
<td>unprotected api</td>
</tr>
<tr>
<td>LifeLabs</td>
<td>2019</td>
<td>15,000,000</td>
<td>healthcare</td>
<td>hacked</td>
</tr>
<tr>
<td>Ministry of Health (Singapore)</td>
<td>2019</td>
<td>14,200</td>
<td>healthcare</td>
<td>poor security/inside job</td>
</tr>
<tr>
<td>Mobile TeleSystems (MTS)</td>
<td>2019</td>
<td>100,000,000</td>
<td>telecommunications</td>
<td>misconfiguration/poor security</td>
</tr>
<tr>
<td>Quest Diagnostics</td>
<td>2019</td>
<td>11,900,000</td>
<td>Clinical Laboratory</td>
<td>poor security</td>
</tr>
<tr>
<td>StockX</td>
<td>2019</td>
<td>6,800,000</td>
<td>retail</td>
<td>hacked</td>
</tr>
</tbody>
</table>
Hacks in 2018

*Facebook - Cambridge Analytica data scandal: 87 million user profiles were disclosed*

- Various political organizations used information from Cambridge Analytica to attempt to influence public opinion:
  - 2015 and 2016 campaigns of United States politicians Donald Trump and Ted Cruz
  - 2016 Brexit (British exit from the European Union) vote
  - 2018 Mexican general election, 2018 for Institutional Revolutionary Party
- Successors: a company run by former officials at Cambridge Analytica, Data Propria, has been quietly working for President Donald Trump’s 2020 re-election effort
Hacks in 2018 (cont.)

Under Armour: a data breach of 150 million accounts, with compromised data consisting of user names, the users’ e-mail addresses and hashed passwords.

Intel x86 microprocessors hardware vulnerabilities: Meltdown and Spectre.
Hacks in 2018 (cont.)

- Saks Fifth Avenue / Lord & Taylor, 5 million credit card holders compromised
- British Airways, a data theft of about 380,000 customer records
- US Centres for Medicare & Medicaid Services (CMS), a data breach that exposed files of 75,000 individuals
- SingHealth, 1.5 million personal data compromised
- Quora reported a data breach that affected its 100 million users data
More in Years Prior to 2018 ...


<table>
<thead>
<tr>
<th>Company</th>
<th>Year</th>
<th>Number</th>
<th>Sector</th>
<th>Issue</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cox Communications</td>
<td>2016</td>
<td>40,000</td>
<td>telecoms</td>
<td>hacked</td>
<td>[90]</td>
</tr>
<tr>
<td>Democratic National Committee</td>
<td>2016</td>
<td>19,252</td>
<td>political</td>
<td></td>
<td>[98]</td>
</tr>
<tr>
<td>US Department of Homeland Security</td>
<td>2016</td>
<td>30,000</td>
<td>government</td>
<td>poor security</td>
<td>[98][110]</td>
</tr>
<tr>
<td>EyeWire</td>
<td>2016</td>
<td>unknown</td>
<td>tech</td>
<td>lost / stolen computer</td>
<td>[130]</td>
</tr>
<tr>
<td>Friend Finder Networks</td>
<td>2016</td>
<td>412,214,295</td>
<td>web</td>
<td>poor security / hacked</td>
<td>[144][145]</td>
</tr>
<tr>
<td>Funimation</td>
<td>2016</td>
<td>2,500,000</td>
<td>web</td>
<td>hacked</td>
<td>[146][147]</td>
</tr>
<tr>
<td>Gyft</td>
<td>2016</td>
<td>unknown</td>
<td>web</td>
<td>hacked</td>
<td>[164][165]</td>
</tr>
<tr>
<td>Inuvik hospital</td>
<td>2016</td>
<td>6,700</td>
<td>healthcare</td>
<td>inside job</td>
<td>[188]</td>
</tr>
<tr>
<td>KM.RU</td>
<td>2016</td>
<td>1,500,000</td>
<td>web</td>
<td>hacked</td>
<td>[196]</td>
</tr>
<tr>
<td>Nival Networks</td>
<td>2016</td>
<td>1,500,000</td>
<td>gaming</td>
<td>hacked</td>
<td>[241]</td>
</tr>
<tr>
<td>Ofcom</td>
<td>2016</td>
<td>unknown</td>
<td>telecom</td>
<td>inside job</td>
<td>[244]</td>
</tr>
<tr>
<td>Rosen Hotels</td>
<td>2016</td>
<td>unknown</td>
<td>hotel</td>
<td>hacked</td>
<td>[281]</td>
</tr>
<tr>
<td>Taobao</td>
<td>2016</td>
<td>20,000,000</td>
<td>retail</td>
<td>hacked</td>
<td>[293]</td>
</tr>
<tr>
<td>TaxSlayer.com</td>
<td>2016</td>
<td>8,800</td>
<td>web</td>
<td>hacked</td>
<td>[297][298][299]</td>
</tr>
<tr>
<td>University of California, Berkeley</td>
<td>2016</td>
<td>80,000</td>
<td>academic</td>
<td>hacked</td>
<td>[330]</td>
</tr>
<tr>
<td>University of Central Florida</td>
<td>2016</td>
<td>63,000</td>
<td>academic</td>
<td>hacked</td>
<td>[332]</td>
</tr>
<tr>
<td>Verizon Communications</td>
<td>2016</td>
<td>1,500,000</td>
<td>telecoms</td>
<td>hacked</td>
<td>[344]</td>
</tr>
<tr>
<td>Weebly</td>
<td>2016</td>
<td>43,430,316</td>
<td>web</td>
<td>hacked</td>
<td>[354][355]</td>
</tr>
<tr>
<td>Bell Canada</td>
<td>2017</td>
<td>1,900,000</td>
<td>telecoms</td>
<td>poor security</td>
<td>[46]</td>
</tr>
<tr>
<td>Defense Integrated Data Center (South Korea)</td>
<td>2017</td>
<td>235 GB</td>
<td>military</td>
<td>hacked</td>
<td>[96]</td>
</tr>
<tr>
<td>Deloitte</td>
<td>2017</td>
<td>350 clients emails</td>
<td>consulting, accounting</td>
<td>poor security</td>
<td>[96][97]</td>
</tr>
<tr>
<td>Equifax</td>
<td>2017</td>
<td>163,119,000</td>
<td>financial, credit reporting</td>
<td>poor security</td>
<td>[120][121]</td>
</tr>
<tr>
<td>Grozio Chirurgija</td>
<td>2017</td>
<td>25,000</td>
<td>healthcare</td>
<td>hacked</td>
<td>[159][180][161]</td>
</tr>
<tr>
<td>Heathrow Airport</td>
<td>2017</td>
<td>2.5GB</td>
<td>transport</td>
<td>lost / stolen media</td>
<td>[175][176][177]</td>
</tr>
<tr>
<td>Taringa!</td>
<td>2017</td>
<td>28,722,877</td>
<td>web</td>
<td>hacked</td>
<td>[394]</td>
</tr>
<tr>
<td>Uber</td>
<td>2017</td>
<td>57,000,000</td>
<td>transport</td>
<td>hacked</td>
<td>[320]</td>
</tr>
</tbody>
</table>
What is Cybersecurity?

Security 101
The Definition of Security

• Security: freedom from, or resilience against, potential harm (or other unwanted coercive change) from external forces (wikipedia) – in physical space

• Cybersecurity: the protection of computer systems from theft or damage to their hardware, software or electronic data, as well as from disruption or misdirection of the services they provide – in cyber space
Cybersecurity Objectives: CIA

Main security objectives:
- **Confidentiality**: unauthorized users cannot read information
- **Integrity**: unauthorized users cannot alter information
- **Availability**: the information must be available when it is needed

Other security objectives:
- Authentication and identification
- Access control
- Anonymity
- Non-repudiation: users cannot deny actions
- Privacy
- ...
Confidentiality

• **The concealment of information or resources**
  - Information is not made available or disclosed to unauthorized individuals, entities, or processes
  - E.g., your bank accounts, private photos, etc

• How to achieve confidentiality? Encrypt the data using a secret key, and only the authorized entities can obtain the secret key to decrypt the data
  - Symmetric encryption: AES, DES, 3DES
  - Asymmetric encryption: RSA
Integrity

• Maintaining and assuring the accuracy and completeness of data over its entire lifecycle
  • Data cannot be modified in an unauthorized or undetected manner
  • E.g., your emails, your electronic homework
Do to Ensure Integrity?

• Generate digest and perform integrity checking

digest (SHA1, SHA2)

Alice

Ave

Bob
Availability

• For any information system to serve its purpose, the service/information must be available when it is needed
  • This means the computing systems used to store and process the information, the security controls used to protect it, and the communication channels used to access it must be functioning correctly

• High availability systems aim to remain available at all times
  • Preventing service disruptions due to power outages, hardware failures, and system upgrades
  • Preventing denial-of-service attacks, such as a flood of incoming messages to the target system, essentially forcing it to shut down
Authentication and Identification

• Authentication in physical world: are you really who you claim?
  • Confirm the identity of a person by validating his/her identity document (e.g., driver license, passport, student ID card)

• Authentication in computers:
  • Confirm whether a person is the owner of a smartphone
  • Confirm whether a person is a user of online banking
  • Confirm whether a website is authentic
How to Do Authentication?

• Four general means for authenticating user's identity
  • Something the individual knows
    • Passwords
  • Something the individual possesses, a *token*
    • Memory card, smart card
  • Something the individual is
    • Fingerprint, iris, retina, face
  • Something the individual does (behavior pattern)
    • Typing rhythm, gait, and voice
How to Do Authentication (cont.)?

- Multi-factor authentication (MFA) – used in our own IT systems in MTU
Access Control

• Access control in physical world: the selective restriction of access to a place. It is a matter of who, where, and when.
  – Historically, this was partially accomplished through mechanical keys and locks

• Access control in computers: the selective restriction of access to computing resources (who, what, and how)
  – Who: users, programs, processes, etc.
  – What: computing resources like files, memory, I/O ports, etc.
  – How: how the computing resources can be “touched”
How to Do Access Control?

• Encrypting the protected computing resources using secret keys, and only disclose keys to those who are authorized.

• The access control is enforced by systems (operating systems, database management systems, etc.) following permissions.
Why Learning Cybersecurity?
Great Job Market

According to US Bureau of Labor Statistics (BLS), the cybersecurity relating job opportunity will be growing 33% every year, which is much faster than the average.
Protect Your Own Asset

• Reduce the possibility of exposure to potential hacks
  • Malicious code is here and there (malicious java scripts, applets, etc.)
  • Make sure you trust the web sites before you go there (a lot of phishing website)
    • www.google.com is fine, but www.go0gle.com may not
    • Do you want to click the link www.facebook.net, or www.b-of-America.co.cc
Protect Your Own Asset (cont.)

• Reduce the possibility of exposure to potential hacks

• A lot of phishing emails
Security Technology Is Money Sometimes

Bitcoin price

Bitcoin price
Mobile Devices and Flash Memory
Mobile Devices are Turning to Mainstream Computing Devices

Number of smartphone users worldwide from 2014 to 2020 (in billions)

Number of tablet users worldwide from 2013 to 2021 (in billions)
Mobile Devices are Turning to Mainstream Computing Devices (cont.)

Number of connected wearable devices worldwide from 2016 to 2021 (in millions)

- 2016: 325
- 2017*: 453
- 2018*: 593
- 2019*: 722
- 2020*: 835
- 2021*: 929
Mobile Devices are Used for Critical Applications

• Mobile devices are increasingly used to handle sensitive data
  – Online banking
  – Ecommerce
  – Cryptocurrency/stock trading
  – Naked photos
  – A human rights worker collects evidence of atrocities in a region of oppression
  – Etc.

• Security issues in mobile computing devices
  – Confidentiality
  – Integrity
  – Authentication
  – Access control
  – ...
A Common Storage System of a Mobile Device

- **Applications layer**: Files, APPs
- **Mobile file system layer**: EXT4, EXT3, EXT2, etc. Implement system calls like open, read, write, etc.
- **Block device layer**: Manage the mappings between the applications’ view and the block device’s view
- **Flash memory layer**: eMMC Chip, MMC Card
Flash Memory Security Research
NAND Flash is Usually Used as Storage Media

- NAND flash
  - USB sticks
  - Solid state drives (SSD)
  - SD/miniSD/microSD/eMMC

**Diagram:**
- **Block** (e.g., 128KB)
- **Page** (e.g., 4KB)

NAND flash is Usually Used as Storage Media.
Special Characteristics of NAND Flash

- **Update unfriendly**
  - Over-writing a page requires first erasing the entire block
  - Write is performed in pages (e.g., 4KB), but erase is performed in blocks (e.g., 128KB)

- Over-write may cause significant write amplification
Special Characteristics of NAND Flash (cont.)

• Support a finite number of program-erase (P/E) cycles
  • Each flash block can only be programmed/erased for a limited number of times (e.g., 10K)
  • Data should be placed evenly across flash (wear leveling)
How to Manage NAND Flash

• Flash-specific file systems, which can handle the special characteristics of NDND flash
  • YAFFS/YAFFS2, UBIFS, F2FS, JFFS/JFFS2
  • Less popular
How to Manage NAND Flash (cont.)

• Flash translation layer (FTL) – a piece of flash firmware embedded into the flash storage device, which can handle the special characteristics of NAND flash and emulate the flash storage as a regular block device (most popular)
  • SSD
  • USB
  • SD
Flash Translation Layer (FTL)

- FTL usually provides the following functionality:
  - Address translation
  - Garbage collection
  - Wear leveling
  - Bad block management
Flash Translation Layer (cont.)

• Address translation
  • Translate address between block addresses and flash memory addresses
  • Need to keep track of mappings between Logical Block Address (LBA) and Physical Block Address (PBA)

The flash translation layer (FTL) should maintain a mapping table:

<table>
<thead>
<tr>
<th>Block device location</th>
<th>Flash location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector 1</td>
<td>(2,3)</td>
</tr>
<tr>
<td>Sector 2</td>
<td>(n,1)</td>
</tr>
</tbody>
</table>
Flash Translation Layer (cont.)

• Garbage collection
  • Flash memory is update unfriendly
  • Not prefer in-place update, but prefer out-of-place update
  • The blocks storing obsolete data should be reclaimed periodically by garbage collection
Flash Translation Layer (cont.)

• Wear leveling
  • Each flash block can be programmed/erased for a limited number of times
  • Distribute writes evenly across the flash to prolong its lifetime
Flash Translation Layer (cont.)

• Bad block management
  • Regardless how good is the wear leveling, some flash blocks will eventually turn “bad” and cannot reliably store data
  • Bad block management is to manage these bad blocks
A Testbed for Flash Memory Security Research

• We have a flash memory testbed in MTU Security and Privacy (SnP) Lab. The lab is located in Rekhi 318

• The testbed includes:
  • Open-source flash firmware: OpenNFM
    • Implement flash translation layer (FTL) which is used to manage raw NAND flash, and provide a block access interface to upper layer
  • Embedded development environment: IAR Embedded Workbench
  • Electronic board: LPC-H3131
A Demo of Flash Memory Testbed

• A demo by Niusen Chen
  • Cross-compile opensource flash firmware OpenNFM using : IAR Embedded Workbench

• Flash the binary to the electronic board LPC-H3131

• Use the electronic board as a USB device (YOU CAN MAKE YOUR OWN USB DEVICE NOW)

• Test throughput using benchmark tool fio
Opensource Flash Firmware OpenNFM
OpenNFM - MTD

• MTD: built on top of raw flash, and mainly provides three uniform APIs to allow the UBI to read, write and erase raw flash

  • MTD_Read(PEB index, offset, &data): read data from a PEB page identified by PEB index and offset

  • MTD_Write(PEB index, offset, data): write data to a PEB page identified by PEB index and offset

  • MTD_Erase(PEB index): erase the PEB identified by PEB index
OpenNFM - UBI

- UBI: built on top of MTD, and uses the APIs provided by MTD to read/write PEB pages or erase PEB blocks. Implement wear leveling, garbage collection, bad block management

  - UBI_Read(LEB index, offset, &data): read data from an LEB page identified by LEB index and offset

  - UBI_Write(LEB index, offset, data): write data to an LEB page identified by LEB index and offset

  - UBI_Erase(LEB index): erase an LEB identified by LEB index, which will cause an erasure over the corresponding PEB
OpenNFM - FTL

• FTL: build on top of UBI, and use the APIs provided by UBI to read/write LEB pages or erase LEBs. Implement address translation

  • FTL_Read(block_address, &data)

  • FTL_Write(block_address, data)
Hands-on Task 1 – Building A Flash Storage Device by Porting Open-source flash firmware to An Electronic Board

- Get familiar with the embedded development environment
- Play with the source code of flash firmware OpenNFM
- Port the OpenNFM to an electronic development board LPC-H3131

- Two students in a group
- Each group will be provided with a desktop and an electronic development board (with cables)
- See the webpage [https://snp.cs.mtu.edu/outreach/wics2022.html](https://snp.cs.mtu.edu/outreach/wics2022.html) for the manual and video tutorial
Conventional full disk encryption (FDE) is incorporated on the block layer to transparently ensure data confidentiality

- BitLocker (Windows)
- FileVault (MAC OS X)
- Android FDE
- TrueCrypt/VeraCrypt

We will move the FDE from the block layer downwards to the flash memory layer (implement an symmetric encryption algorithm you like)

- Write a flash page, encrypt it
- Read a flash page, decrypt it using the same key
- Encryption/Decryption is completely transparent to users
- Test the throughput using benchmark tool fio, compare the throughput with original OpenNFM firmware