Friedrich-Alexander-Universität
Faculty of Engineering

Cybercrime and
Forensic Computing
Research Training Group 2475

Flashback: Extending a Study of Flash Sanitization Practices

Janine Schneider, Friedrich-Alexander- Universität Erlangen-Nürnberg

Joint work with Aya Fukami[a], Immanuel Lautner[b], Denise Moussa[b], Julian Wolf[b], Nicole Scheler[b], Dominic Deuber[b], Felix Freiling[b], Jaap Haasnoot[c], Hans Henseler[c], Simon Malik[d], Holger Morgenstern[d] and Martin Westman[e]

Background

- In 2018 Martin Westman reported that he had found non-trivial data on new USB drives. It has been speculated that Westman's findings were due to the reuse of NAND flash chips in USB devices. [1]
- Therefore, in 2021 we acquired **650 low-cost USB drives** and **analyzed 614** of them, in order to **assess the risk of acquiring evidence on newly purchased USB drives** originating from NAND flash chip recycling. [2]
- We extended the study by acquiring **another 600 low-cost USB drives** and **459 branded high-cost USB drives.** We **analyzed 589 low-cost and 435 high-cost drives.**

Study Execution

- Decentralized low-cost USB drive acquisition in batches via Alibaba (ordered by cost).
- Centralized high-cost USB drive acquisition in small batches via several distinguished German online shops (ordered by brand).
- Measurements:
  - USB drive manufacturer
  - Chip manufacturer
  - Physical appearance
  - Chip type (raw NAND or eMMC)
  - NAND technology
  - Capacity in GB
  - Cost
- Analysis steps:
  - Label drive
  - Take forensic 1:1 image
  - Gather measurements
  - Carve for data (scalpel, foremost)
  - Calculate entropy (ent, binwalk)
  - Disassemble drive
  - Chip-off

	Low-cost				High-cost	
	FAU	HSL	HSAS	Total	FAU	Total
Drives	516	134	600	1,250	459	1,709
Analyzed	489	133	589	1,211	435	1,646
Data found	61	14	1	76	0	76
Visual Inspection	415	0	555	970	305	1,275
Entropy	479	119	89	687	453	1,140
Chip-off	8	0	5	13	16	29

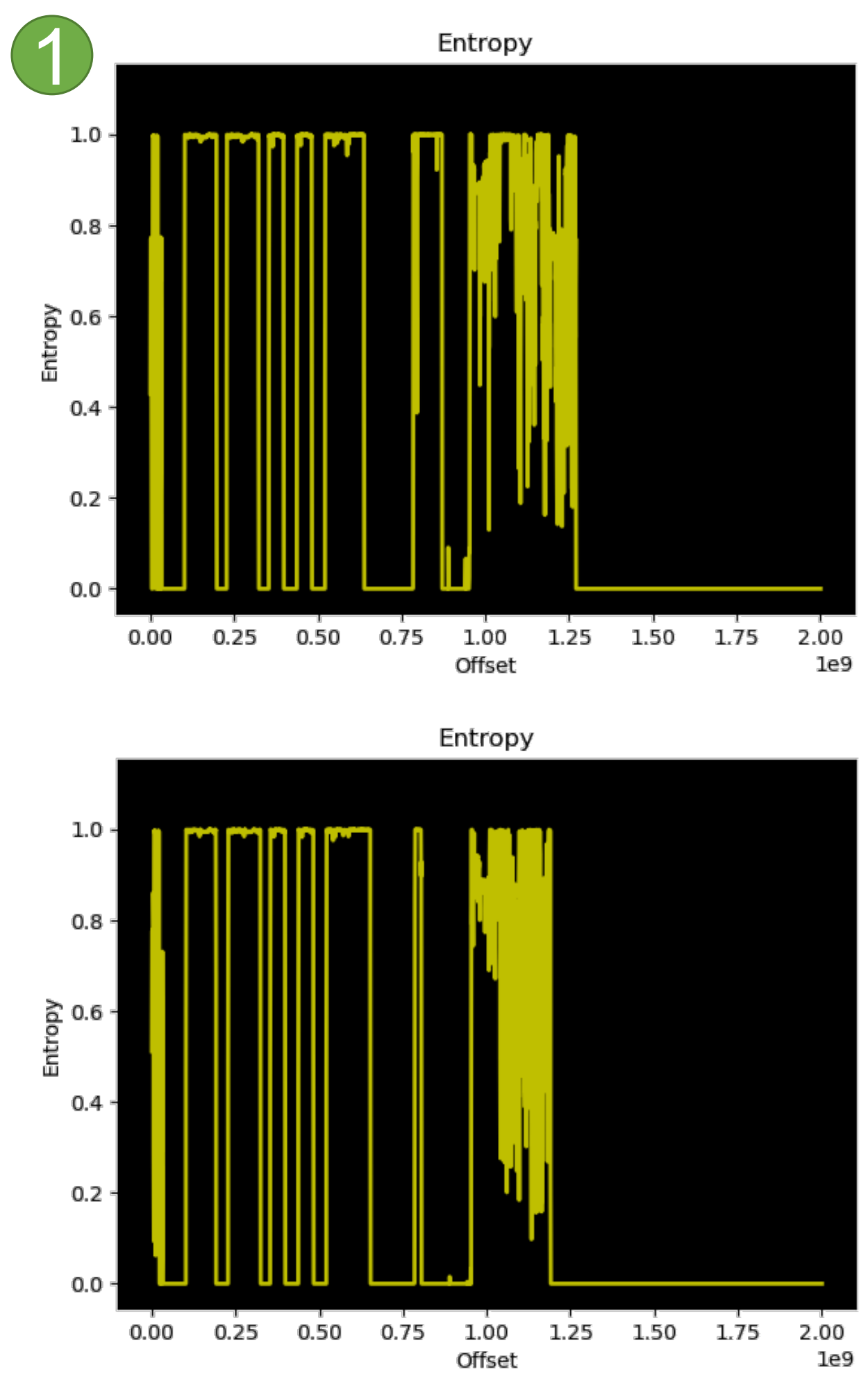
High-Cost Device Results

- **None of the analyzed high-cost USB drives contained non-trivial data** originating from chip recycling.
- Some were shipped pre-formatted and contained pictures of the brand icon.
- The visual inspection revealed **irregular engravings, markings and epoxy.** 5
- The entropy analysis showed medium to high entropy for some devices. Some of them contained test files, some showed unexplainable entropy peaks and patterns, some contained random data and/or 0xff. It could be that the data originates from functional tests of the chip manufacturers.

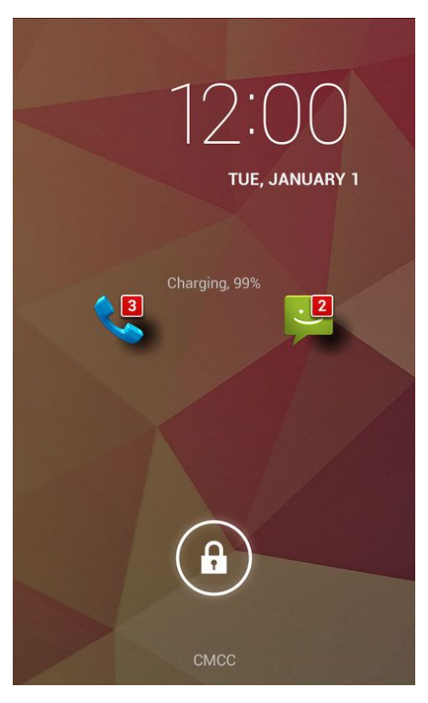
Low-Cost Device Results

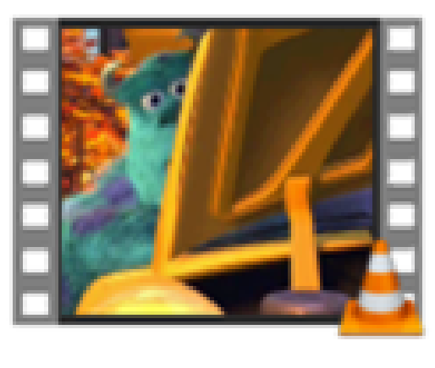
- **76 USB drives contained non-trivial data.**
- 2 USB drives contained an active FAT32 filesystem containing deleted private pictures (probably originating from testing the device).
- On the remaining 74 drives we found **media data, maps, OS data, documents, speech recordings and source code.**
- The data could be assigned to **Android, Chrome and Linux OS, Printers, Navigation Systems, Smart TVs** and other devices. 2
- The visual inspection revealed **various impurities, remnants, scratches, irregular stamps and engravings.** 4
- Some USB drives contained **cut mini-SD cards or shortened chips**. Cutting or shortening the chips is a known procedure to disconnect the internal connection between the NAND flash and the controller if the controller fails the functional test. 3
- The entropy analysis showed high entropies for some USB drives not containing non-trivial data (probably overwritten or encrypted).
- The entropy analysis for USB drives containing non-trivial data showed **resembling patterns.** 1
- Through chip-off the eMMC´s health reports could be read which revealed that these chips had already performed **hundreds of erase cycles.**

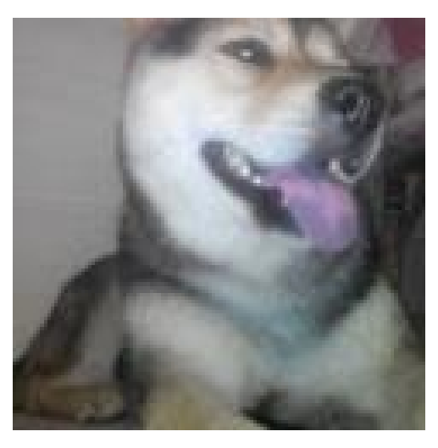
1



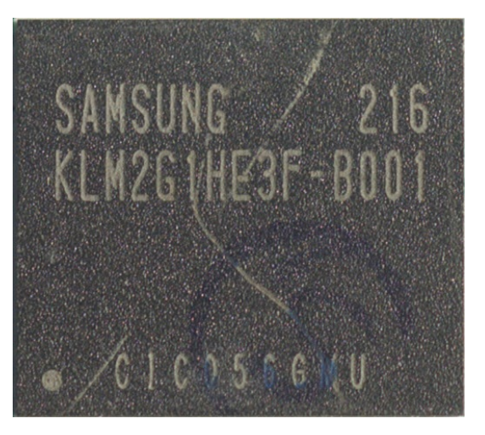
2

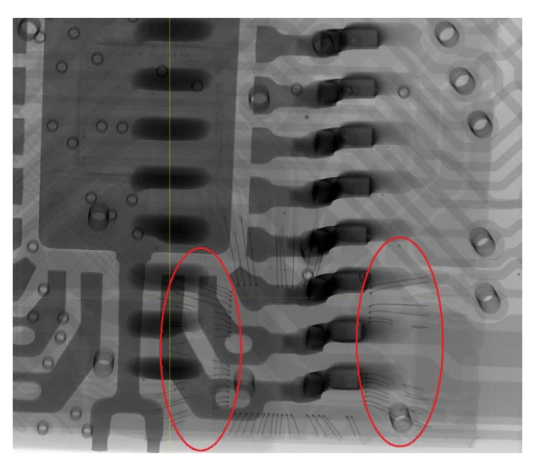


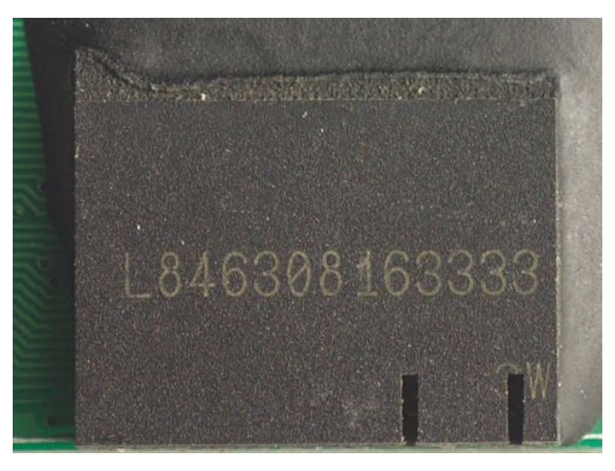




4





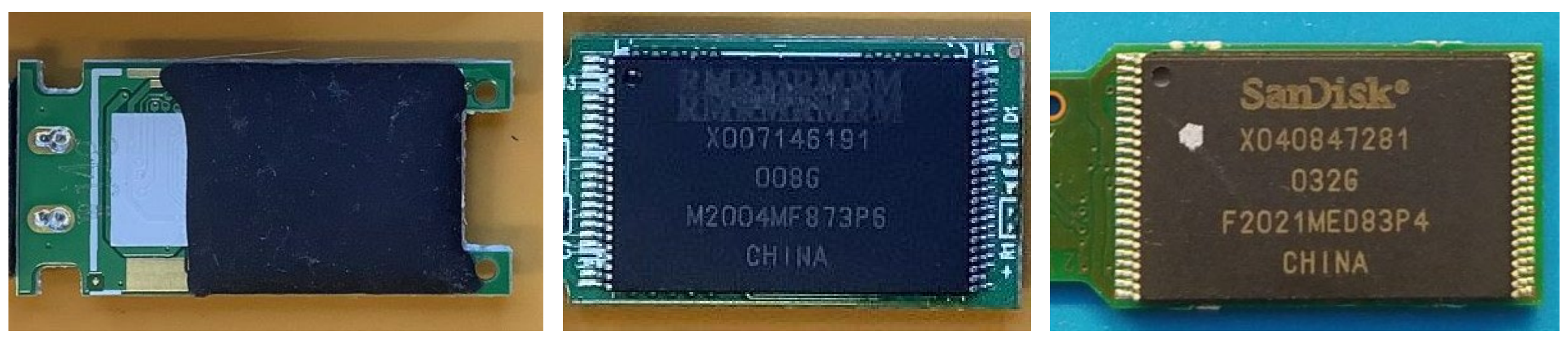


3

Conclusion

- We found a probability of 6% of finding data on cheap but new USB drives.
- This probability hardly depends on the supplier of the drives, as we observed low-cost suppliers with a nearly 100% probability of finding non-trivial data originating from chip recycling.
- The probability of finding non-trivial data on new branded higher-priced USB drives is approximate 0%.

5



References

1. Martin Westman, Where did that incriminating evidence come from?, DFRWS EU (2018)
2. Janine Schneider et al., In Search of Lost Data: A Study of Flash Sanitization Practices, DFRWS EU (2021)
3. Aya Fukami, Sasha Sheremetov, Francesco Regazzoni, Zeno Geradts and Cees De Laat, Experimental Evaluation of eMMC Data Recovery, IEEE Transactions on Information Forensics and Security (2022)

a Netherlands Forensic Institute, The Hague, The Netherlands
b Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
c Leiden University of Applied Sciences, Leiden, The Netherlands
d Albstadt-Sigmaringen University, Albstadt, Germany
e Micro Systemation (MSAB), Stockholm, Sweden

Contact: Janine Schneider <janine.schneider@fau.de>