Bias in the TRNG of the Mifare DesFire EV1, a CC EAL 4+ RFID Card, and what went wrong

Darren Hurley-Smith
Julio Hernandez-Castro

19/1/2017

Outline

Introduction

University of Kent | Computing

- We are researching potential vulnerabilities in the Mifare DESFire EV1, and others: EV2, FeliCa, LEGIC.
- The DESFire EV1 is a Common Criteria EAL4+ certified smart card used in transport, Univ. student/staff ID, fare payment and other micro-payments, around the world [1]
- It is also used in loyalty schemes and access control applications
- Transport for London (TfL) issued approximately 8 million DESFire EV1-based cards in the 2015/2016 period [2]

Motivation: Background

University of Kent | Computing

- A weak PRNG and cryptographic algorithm (CRYPTO-1) contributed to dismantling the Mifare Classic [3–6]
- Let's do it again!
- Also, they continue to use **security by obscurity**, not publicising anything about the TRNG design they use
- Let's prove again that's not a good idea!

Motivation: Rationale

University of Kent | Computing

- As a cash-value bearing card, the DESFire EV1 has a monetary value to criminals
- The Mifare DESFire EV1 has been successfully emulated [7], and its power characteristics have been analysed in depth [8]
- This card has proven resilient to side-channel attacks (SCA), by implementing hardware countermeasures [9]
- We believed that an in depth evaluation of the DESFire EV1's TRNG could be interesting

Methodology: Data collection

University of Kent | Computing

- 64 MB of data was retrieved from three DESFire EV1 and one EV2 using an ACR122U reader
  - Data collection took an average of 12 days per card to gather 4 million AES-128 encrypted values
  - Each nonce (16-bytes long) was extracted from a different authentication session
- The data was acquired from the protocol used to secure card PICC and Application read/write functions
- The values were decrypted using a default AES-128 key (initialised to zero) before analysis

Methodology: Lab Set-up

(a) Laptop with reader and DESFire EV1 cards

(b) ACR122U reader with two of three Mifare DESFire EV1 cards

Figure 1: Experimental set-up used to collect TRNG data

- Toshiba Laptop Specification: i7 processor, 8GB RAM
- Reader: ACR122U (CCID), Scripts: Python 2.7 and Bash

Methodology: Randomness Tests

The collected data was subjected to three randomness test batteries:

- **The NIST Statistical Test Suite v2.1.2**
- Dieharder
- ENT

Results: NIST STS 2.1.2

Table 1: Mifare DESFire EV1 ENT results for 64MB of TRNG output

	Passed Tests
Card 1	198/200
Card 2	200/200
Card 3	197/200

- All cards passed the NIST STS 2.1.2 battery within acceptable parameters (greater than 193)
- Cards 1 and 3 perform poorly in one of the non-overlapping template tests, but passed all other tests
  - This seems to be statistically insignificant
- Card 2 passes all tests with no weak results

Results: Dieharder

University of Kent | Computing

Table 2: Diehard results for 64MB of TRNG Output

Diehard Test		EV1 Card 1	EV1 Card 2	EV1 Card 3
	t-samples	p-values	p-values	p-values
Birthday Spacings	default	0.18194520	0.61105583	0.78263630
Overlapping Permutations	125,000	0.38044164	0.58693289	0.44201308
6x8 Binary Rank	25,000	0.31311490	0.32387215	0.66137580
Bitstream	default	0.97724174	0.18743536	0.45949716
Count the 1's (stream)	default	0.17108396	0.74984724	0.87214241
Count the 1's (byte)	default	0.86481241	0.92578024	0.00000000
Parking Lot	default	0.18078043	0.24200626	0.38128677
Minimum Distance (2d sphere)	default	0.76328000	0.95091635	0.34980807
3d sphere (minimum distance)	default	0.23871272	0.20826216	0.39340851
Squeeze	default	0.62598919	0.08843989	0.00026939
Runs	default	0.63756836	0.80941394	0.04870531
Craps	20,000	0.54077256	0.92769962	0.91803037

Results: ENT

Table 3: Mifare DESFire EV1 ENT results for 64MB of TRNG output

ENT	EV1 Card 1	EV1 Card 2	EV1 Card 3	Optimal/Expected
Entropy	7.999969	7.999989	7.999972	8
Optimal Compress.	0	0	0	0
chi-square	2709.10	973.07	2470.32	255
Arith. Mean	127.492921	127.500582	127.5006	127.5
Monte Carlo π est.	3.14167	3.142019	3.141909196	3.14159
S. Correlation	0.000008	0.000045	0.000093	0.0

- Both cards demonstrate very poor performance on the chi-square test
- This indicates that there is a strong bias in the distribution of byte values throughout both data samples

Results: ENT 2 (1MB Sample)

University of Kent | Computing

Table 4: Mifare DESFire EV1 ENT results for 1MB of TRNG output

ENT	EV1 Card 1	EV1 Card 2	EV1 Card 3	Optimal/Expected
Entropy	7.999780	7.999820	7.999786	8
Optimal Compress.	0	0	0	0
chi-square	305.47 (1.65%)	249 (59.41%)	297.03 (3.62%)	255
Arith. Mean	127.6015	127.5626	127.5082	127.5
Monte Carlo π est.	3.13620558	3.140892564	3.140388562	3.14159
S. Correlation	-0.000068	0.001339	-0.001751	0.0

- All cards perform better on this test with smaller samples and have p-values greater than 0.01 for the chi-square test
- This indicates that the bias may be missed if tests are not performed on a large enough sample of TRNG data (greater than 1 MB in the case of ENT)

Results: ENT 2 (5KB Sample)

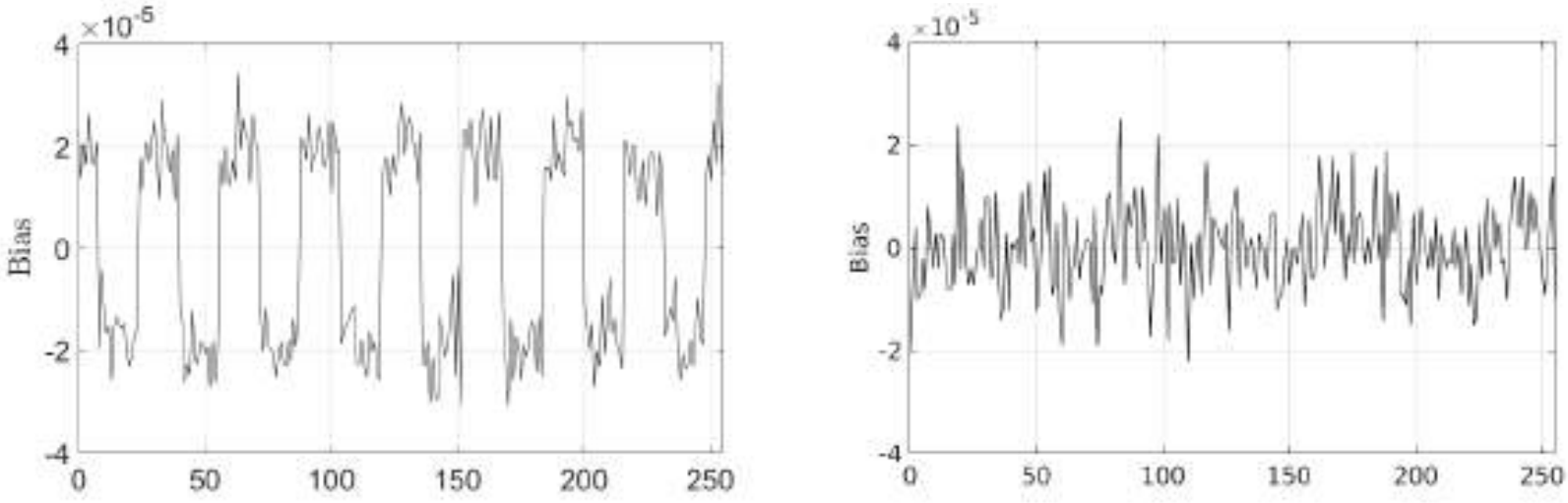
University of Kent | Computing

Table 5: Mifare DESFire EV1 ENT results for 5KB of TRNG output

ENT	EV1 Card 1	EV1 Card 2	EV1 Card 3	Optimal/Expected
Entropy	7.999635	7.999640	7.999641	8
Optimal Compress.	0	0	0	0
chi-square	253.55 (51.3%)	249.26 (58.96%)	249.03 (59.36%)	255
Arith. Mean	127.6015	127.6324	127.4534	127.5
Monte Carlo π est.	3.13744549	3.145452582	3.140388562	3.14159
S. Correlation	-0.000579	0.001990	-0.001204	0.0

- At 5KB, all tests pass the chi-square test with p-values greater than 0.01 and within acceptable bounds of the expected value
- Deviations start to suggest themselves (but not totally clear) in the chi-square results for samples larger than 7.5KB.

Analysis: Bias



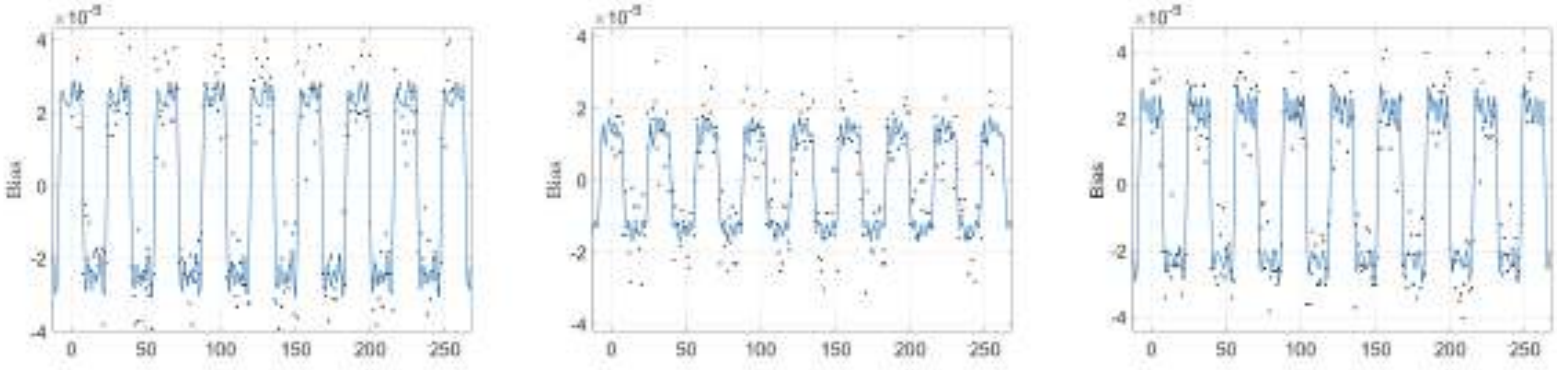
(a) Mifare DESFire EV1 mean bias

(b) Random Data

Figure 2: Bias of two 64MB datasets

- Only (a) shows a clear non-random trend
- repetitive pattern, clear cycles, almost no values close to zero...

Analysis: Fourier Analysis of the Bias



(a) EV1 Card 1 (b) EV1 Card 2 (c) EV1 Card 3

Figure 3: Fourier series for the biases from three 64MB TRNG samples

- All cards demonstrate a regular period of 32 biased values
- Exactly half of the possible byte values occur more frequently

Initial Findings

University of Kent | Computing

- Previously, evaluators have used AIS-31 (CC)
- and researchers employed NIST, Dieharder, chi-square, and other uniformity tests[1]
- but these are not always sufficient to find the bias in the EV1 TRNG
- Chi-square tests performed in the literature focus on bit-level analysis, but the bias is only apparent at the byte level in this TRNG

[1] Private communication.

Further Findings

University of Kent | Computing

- We experimentally found that bits 4 and 5 of every byte seemed to frequently take the same value
- We built a model with different probabilities of this occurring, and found heuristically that the one with a probability of

$$p(b_4 = b_5) = 1/2 - \epsilon \text{ with } \epsilon \sim 0.0004325$$

performed the best in approximating the bias of the three EV1 cards

- After these analysis, we can explain up to $R^2 \approx 0.8121$ (average of 77.96%) of the observed bias for EV1

Further Findings

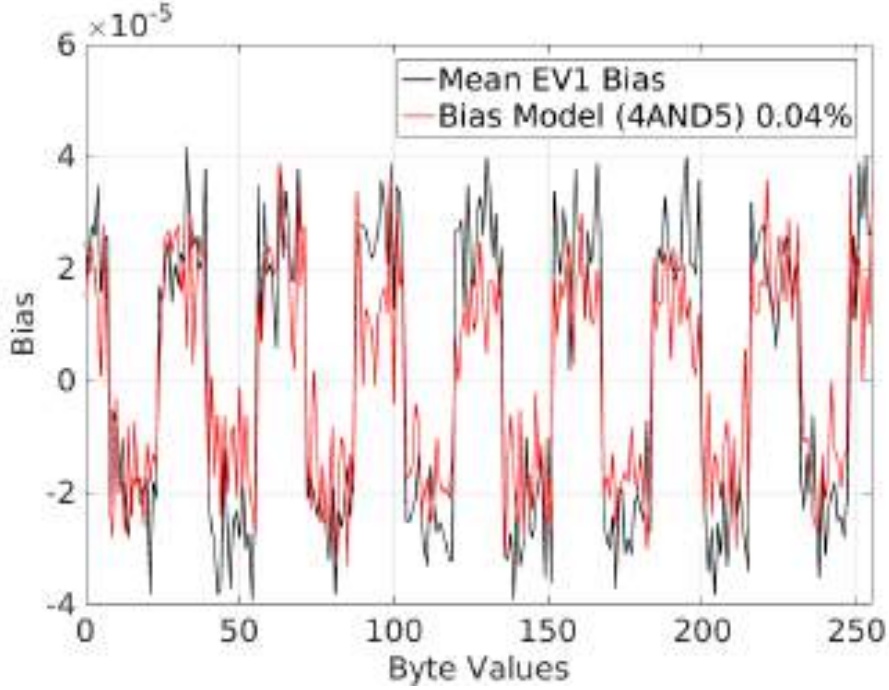
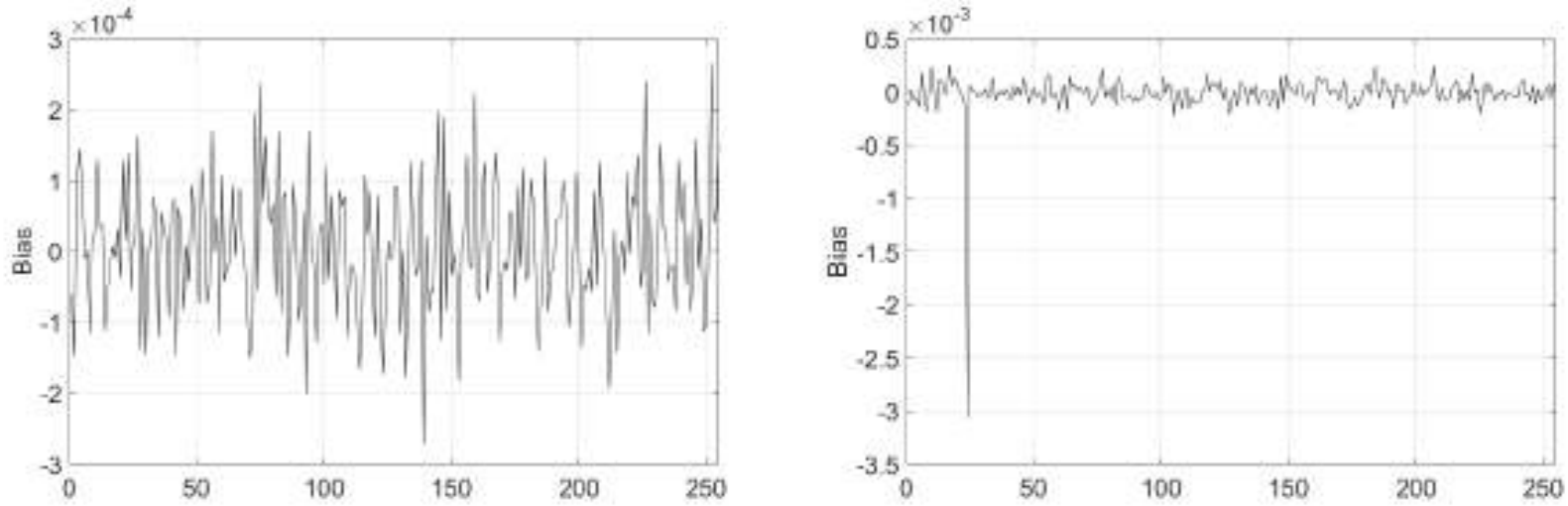


Table 6: Coefficient of determination (R^2)

	R^2	Adjusted R^2
Card 1	0.7981	0.7836
Card 2	0.7288	0.7094
Card 3	0.8121	0.7987
urandom	0.0479	-0.0201

Analysis: New Randomness Test: The Bitmask Test

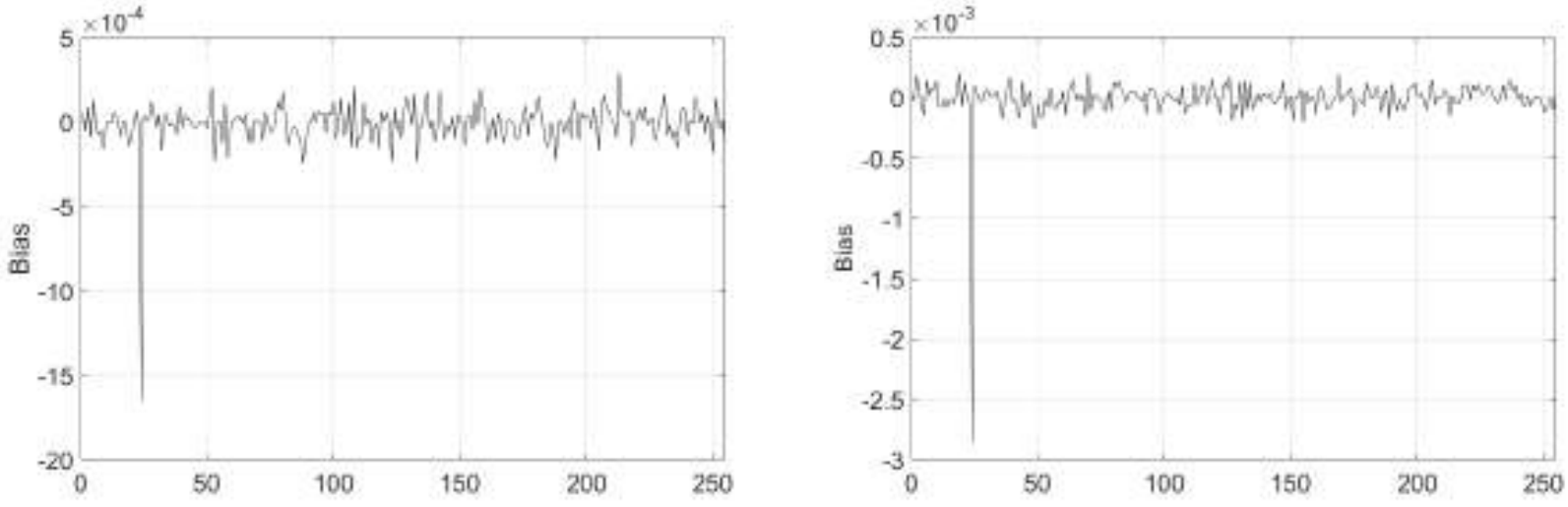


(a) Random Source (b) EV1 Card 1

Figure 4: Graphs showing the bias of masks applied over 64MB samples

- Basic idea is to apply linear cryptanalysis to the raw random data **r**
- We look for $max_{m \in (1..255)} | \sum m \cdot r - r/2 |$
- Maximum bias for m=24 (00011000), consistently

Analysis: Bit-Mask Test Results



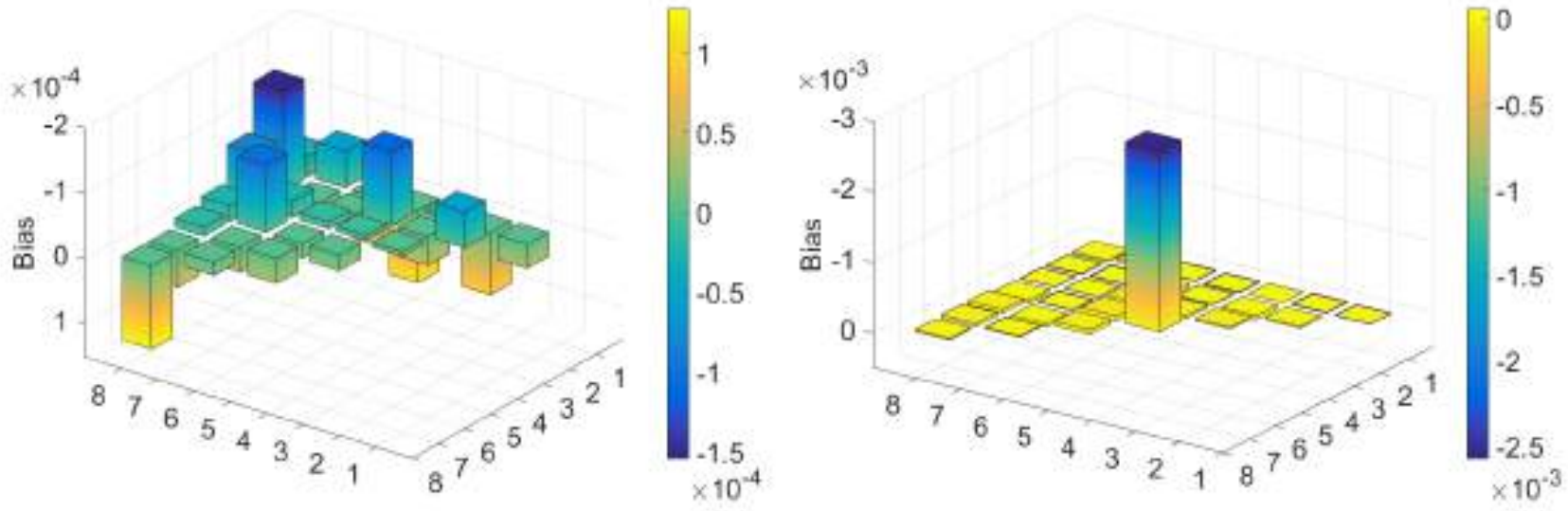
(a) EV1 Card 2

(b) EV1 Card 3

Figure 5: Graphs showing the bias for all masks applied over 64MB samples

- This trend holds across all three EV1 64MB data sets

Analysis: Byte-level Analysis



(a) Random Source (b) Mean EV1 Bias

Figure 6: Visualisation of the bias of bit-adjacency for all bytes (64M)

- (a) shows the same visualisation as (b) over random data
- (b) shows that bits 4 and 5 have a predisposition to sharing values

Conclusion

University of Kent | Computing

- We have conducted a study of the Mifare DESFire EV1's 'true' random number generator
  - Clear & consistent biases have been found in the data
- We have responsibly disclosed our findings to NXP
  - They have responded, confirming our findings
  - They have a team "looking into the root cause"
- No practical attacks have been identified at this point
  - But hopefully future attacks will build on these findings
- We have observed that some of the best known tests do not detect this flaw, PRNG/TRNG evaluation is tricky!
- In particular, we have shown some of the limitations of the current CC certification process and would recommend major changes on it
  - And we have worked with average case, not worst case scenarios!

Future Work

University of Kent | Computing

- Continue to collect data from more Mifare DESFire EV1 (check manufacturing issues) and EV2 cards
- Expand the tested cards to include other Common Criteria EAL 4/5+ RFID smart cards:
  - Felica
  - Legic
- Testing collected data with other test batteries (e.g. TestU01)
- Testing under variable environmental conditions (extreme temperatures, etc.)
- Developing hardware model that explains the observed bias (help wanted!)
- Start an ambitious project to analyse current randomness test suites and study their test's independence and sensibility to come up with a new one to which only new independent test will be added.

Acknowledgements

- This work was funded by InnovateUK as part of the authenticatedSelf project, under reference number 102050.

- We would like to thank ECOST - Cryptacus Action for their valuable and insightful discussion of this work.
- We would also like to thank NXP Semiconductors Ltd. for their timely and professional communication following the responsible disclosure of our findings.

References

1. NXP Semiconductors. *MIFARE DESFire EV1 4K: MIFARE DESFire EV1 contactless multi-application IC*. Retrieved 15:45 05/09/2016 from: http://www.nxp.com/products/identification-and-security/mifare-ics/mifare-desfire/.
2. Transport for London. *Adult Oyster Cards Issued 2015/16*. http://content.tfl.gov.uk/oyster-card-sales.pdf, 2016.
3. Gerhard de Koning Gans, Jaap-Henk Hoepman, and Flavio D Garcia. A practical attack on the mifare classic. In *Int. Conference on Smart Card Research and Advanced Applications*, pages 267–282, 2008.
4. Mohamad Merhi, Julio Hernandez-Castro, and Pedro Peris-Lopez. Studying the prng of a low-cost rfid tag. In *RFID-Tech. and Applications (RFID-TA), 2011 IEEE Int. Conference on*, pages 381–385.
5. Flavio D Garcia, Peter Van Rossum, Roel Verdult, and Ronny Schreur. Wirelessly pickpocketing a mifare classic card. In *30th IEEE Symposium on Security & Privacy*, pages 3–15, 2009.
6. Yi-Hao Chiu, Wei-Chih Hong, Li-Ping Chou, Jintai Ding, Bo-Yin Yang, and Chen-Mou Cheng. A practical attack on patched mifare classic. In assic, editor, *Int. Conference on Information Security and Cryptology*, pages 150–164. Springeron, 2013.
7. Timo Kasper, Ingo Von Maurich, David Oswald, and Christof Paar. Chameleon: A versatile emulator for contactless smartcards. In *International Conference on Information Security and Cryptology*, pages 189–206. Springer, 2010.
8. Timo Kasper, David Oswald, and Christof Paar. Side-channel analysis of cryptographic rfids with analog demodulation. In *International Workshop on Radio Frequency Identification: Security and Privacy Issues*, pages 61–77. Springer, 2011.
9. Timo Kasper, Ingo von Maurich, David Oswald, and Christof Paar. Cloning cryptographic rfid cards for 25usd. In *5th Benelux workshop on information and system security. Nijmegen, Netherlands*, 2010.

Questions?