Cryptanalysis of Ascon - An Information Theoretic Perspective - A Position Paper

Nicolas T. Courtois, Florian Caullery, Fred Amiel, William Whyte
QCT Security Architecture, Sophia Antipolis, Qualcomm France S.A.R.L.
Technical Standards, Qualcomm Technologies, MA, US
Agenda

- New encryption standard!
- A novel approach to analyze the security of Ascon
- Strong and weak S-boxes: do we need two theories?
- How attacks can be mapped to basic properties
- Prediction of attacks and undesirable properties
- Contemplating the gap and combinatorial explosion
- Open problems
Background - Facts

Feb 2023  
NIST selects ASCON to become a new encryption standard expected to be in use for many decades to come.

June 2023  
NIST hosting the 6th LWC Workshop: NIST is soliciting research and discussion papers, surveys, presentations, panel proposals, case studies related to ASCON including Security results on the Ascon family + a call for public comments. Submission deadline = 1st May 2023.

• Through this presentation, our focus is to engage with NIST to support the community effort to develop the best possible encryption standard. We need to optimize the security and yet minimize HW implementation cost.

• A sensible analysis of security of Ascon should be:
  - **Forward-looking:** we cannot contemplate just some attacks already studied.
  - **Robust:** we should not just look for some rare and exceptional events (best case). We need methods to study understand what happens on average. We claim that there exists a ROBUST transparent way for evaluating a security of a cipher seen as a communications channel trying to maximize the “channel capacity”.
  - **Relevant:** Several already known attacks CAN be modelled in terms of intersections of spaces of some “undesirable properties”.

• **Methodology:** “transforming a constant into a variable”
  - replacing the S-box by several candidates, weak or strong,
  - showing how the attacks scale and showing that their existence can be reliably predicted from the following principles:
    - conditional entropy, mutual information and, discrete combinatorial events weighted by probabilities, which exist in small finite numbers because the S-box is tiny.

Ascon NIST specification

![Diagram of Ascon NIST specification](image)
Qualcomm brought you foundational communications technologies.

Can information theory help cryptographers to design better ciphers?
Analysis of Ascon
Modelling Ascon as a Communications Channel

For 9 years Ascon was studied and seems very secure. All because of “strong diffusion”. Any simple perturbation expand very quickly. Game over = no hope to attack Ascon???

It should be critical to consider attacks that AGGREGATE input perturbations.

---

<table>
<thead>
<tr>
<th>Cipher</th>
<th>( r_e )</th>
<th>( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascon [Dob+16]</td>
<td>3</td>
<td>298</td>
</tr>
<tr>
<td>Gift [Ban+17]</td>
<td>3</td>
<td>60</td>
</tr>
<tr>
<td>Keccak [Ber+11]</td>
<td>2</td>
<td>546</td>
</tr>
<tr>
<td>Present [Bog+07]</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>Pride [Alb+14]</td>
<td>2</td>
<td>31</td>
</tr>
<tr>
<td>Qarma [Ava17]*</td>
<td>2</td>
<td>36</td>
</tr>
</tbody>
</table>

Grassi, Rechberger and Rønjom, 2016, Subspace Trail Cryptanalysis

---

**Table 2: Undisturbed Bits of Ascon’s S-box.**

<table>
<thead>
<tr>
<th>Input Difference</th>
<th>Output Difference</th>
<th>Input Difference</th>
<th>Output Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>00001</td>
<td>?1??</td>
<td>10000</td>
<td>?0??</td>
</tr>
<tr>
<td>00010</td>
<td>1??1</td>
<td>10001</td>
<td>10??1</td>
</tr>
<tr>
<td>00011</td>
<td>??0?</td>
<td>10011</td>
<td>0??0</td>
</tr>
<tr>
<td>00100</td>
<td>??110</td>
<td>10100</td>
<td>0?1??</td>
</tr>
<tr>
<td>00101</td>
<td>1????</td>
<td>10101</td>
<td>?????1</td>
</tr>
<tr>
<td>00110</td>
<td>??????</td>
<td>10110</td>
<td>1????</td>
</tr>
<tr>
<td>00111</td>
<td>0???1?</td>
<td>10111</td>
<td>0???0</td>
</tr>
<tr>
<td>01000</td>
<td>??11?</td>
<td>11000</td>
<td>??1??</td>
</tr>
<tr>
<td>01011</td>
<td>??1?1</td>
<td>11010</td>
<td>??0??</td>
</tr>
<tr>
<td>01100</td>
<td>??00?</td>
<td>11100</td>
<td>??1??</td>
</tr>
<tr>
<td>01110</td>
<td>?0???</td>
<td>11110</td>
<td>?1???</td>
</tr>
<tr>
<td>01111</td>
<td>?1?0?</td>
<td>11111</td>
<td>?0??</td>
</tr>
</tbody>
</table>
Philosophy : Aggregate Perturbations

• Can several perturbations converge somewhat? Attacker does either A or B.

• We need to improve the “channel capacity” to increase information conveyed or the likelihood of detection.
A Known Problem - Analogy with Optics

Not if we have TOO many sources!
Must restrict the input diversity.

Wide Open Aperture

Stopped Down Aperture

a better transmission channel!
Study of **Conditional Entropy and MI**

\[ H(F(x) - F(y) \mid x-y) \]

Prediction ability <=> Mutual Information = \( MI = \)

\[ MI(F(x)-F(y); \ x-y) \]

Input difference \( \xRightarrow{\text{PREDICT}} \) Output difference

\[ \text{should be large} \]

\[ \text{should be small} \]
Ascon S-Box - Proof of Concept

**best =**
- \( \text{Ent}(oD) = 2.00 \) bits when \( x-y=4 \)
- \( \text{Ent}(oD) = 2.00 \) bits when \( x-y=12 \)
- \( \text{Ent}(oD) = 2.00 \) bits when \( x-y=16 \)
- \( \text{Ent}(oD) = 2.00 \) bits when \( x-y=17 \)

**single differences vs. quadruple differences**
- \( \text{Ent}(oD) = 3.69 \) bits when \( x-y \in \{1,3,16,18\} \)
- \( \text{Ent}(oD) = 3.69 \) bits when \( x-y \in \{4,8,20,24\} \)
- \( \text{Ent}(oD) = 3.66 \) bits when \( x-y \in \{5,16,17,21\} \)

we compute the entropy for the output difference

We gain something:
- **av. Output \( \Delta \text{Ent} = 2.00 \) bits**
- **av. Output \( \Delta \text{Ent} = \text{only } 3.69 \) bits instead of 4 bits**
Towards [Difference] Prediction
Prediction approach based on $\text{MI} = \text{Mutual Information}$

Entropy:

$\text{DMI} = \text{Differential Mutual Information}$

For 5 bits we get:

<table>
<thead>
<tr>
<th></th>
<th>Fides($x^{-1})\ APN</th>
<th>RP*</th>
<th>Ascon/Keccak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fides</td>
<td>3.875</td>
<td>3.5</td>
<td>4.10</td>
</tr>
<tr>
<td>Ascon/Keccak</td>
<td>1.125</td>
<td>1.4-1.7</td>
<td>1.90</td>
</tr>
</tbody>
</table>

2x bad

*RP=Random Permutation
**DDT is holographic**

- **Fundamental Observation:**
  - The fact that there are many zeros and some large integers at specific locations are actually the SAME.
  - one cannot happen without the other.

- **Not new - many papers on conflicting security criteria in ciphers and Boolean functions**

  - **What is NEW?** Showing they two properties coincide “EN MASSE” and in probability - they correspond to TWO large precisely measurable percentages of “basic events” which have a large intersection inside an objective information theoretic averaged measure of quality which is simply a SUM weighted by probabilities.

  Overlap between different attacks is a totally objective feature.

---

**“white box security analysis”** => leading to total awareness of basic facts which imply a larger family of attacks
Ascon Relies on Just ONE Tiny NL component

6,8,12 rounds like this:

64x columns of 5 bits

5x lines of 64 bits
Key Problem

Prediction ability?

Claim: Ascon is not very strong in this respect compared to other encryption algorithms... Here is why.
So What?

Imagine that the attacker is trying to break the Ascon hash function by a sophisticated guess then determine attack [see work of Xiaoyun Wang, Marc Stevens, Leo Perrin etc etc]

⇒ when MI is 2x bigger, the amount of information the attacker “already knows” doubles...

We intend to show [very early, just argue] and claim that a large MI has a dramatic impact enabling all of the following:

- all sorts of guess and determine attacks
- truncated differential attacks
- polynomial invariant attacks
- subspace trail attacks
- zero-sum and cube attacks

Claim: a rapid combinatorial explosion.

Really?
Ascon has very few rounds!! Example:

the attacker studies 12 rounds

attacker
128 bits
2/5

A common misconception in cryptanalysis

Is Ascon box good enough?

Claim that... “bad quality” NL mappings are OK if you have a large number of rounds. The problem: Ascon does NOT have a large number of rounds.

Some attacks are such that additional rounds do NOT increase security

1. polynomial invariant attacks and affine space trail attacks ...
2. some differential attacks: very surprising but real...

Composability violation: Proof of concept:
=> Attack works equally well for say 8 and 400 rounds...
Methodology - Example

Example: consider a 5-round truncated differential property on 29 bits of Ascon out of 320.

We can MAP in a precise exact and undeniable way this property seen as an enumeration of discrete cryptographic events to an information-theoretic measure of quality of

A. the direct product of the S-box with itself (parallel application) which has 100% predictable properties in terms of MI and mappings.

B. the linear layer which also has well-defined entropy and mutual information properties w.r.t diffusion.

THEN one can show that our enumeration of events accounts EXACTLY for 43% and 57 % of the MI / entropy in percentage and probability mass for A and B respectively. All attacks comes back to few basic information theoretical facts!

From here we claim that there exists a simple and robust methodology for evaluating the quality of ciphers.

Table 10: Summary of attacks on Ascon.
Are Many Tiny Boxes Toxic?

- 6,8,12 rounds like this:

  **“semi-transparent”**

  ![Semi-transparent diagram](image1)

  ![Semi-transparent diagram](image2)

  ![Semi-transparent diagram](image3)

  **“totally-transparent”**

  ![Totally-transparent diagram](image4)

  ![Totally-transparent diagram](image5)

  ![Totally-transparent diagram](image6)

  Ascon

  unrelated?
Are Many Tiny Boxes Toxic? **YES, without any doubt**

- 6,8,12 rounds like this:

This is a **HUGE** amount of shared information
- An **UNDENIABLE** information-theoretic property
- Active in essentially any of already known attacks on Ascon...

Ascon 64*1.91
122 bits
Are Many Tiny Boxes Toxic? **YES, without any doubt**

- 6,8,12 rounds like this:

  ![Diagram](image)

  Ascon would be only 42 bits total if we used the AES S-box. **122 bits**

  Information-theoretical undeniable cannot be ignored
How high MI implies “Undesirable Properties” or does it?
Academic Background:

The notion of so called “Forbidden Mappings”

**Def.** We call **Sidon-Rodier-Golomb = SRG**\(_2\) mappings all sets of 4 points which map an affine space of dim 2 to an affine space of dim 2 [\(\leftrightarrow\) partial linearity on 4 points]

<table>
<thead>
<tr>
<th>APN AES-like</th>
<th>RP</th>
<th>Ascon/Keccak</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.125</td>
<td>1.57</td>
<td>bad</td>
</tr>
<tr>
<td>0</td>
<td>some exist</td>
<td>bad</td>
</tr>
</tbody>
</table>

Forbidden Mappings =>
MI is Additive and influenced by ALL partial linearity events => 
Prediction Ability

The number of SRG$_2$ or “Forbidden Mappings” is predictable:

<table>
<thead>
<tr>
<th>APN</th>
<th>RP</th>
<th>Ascon/Keccak</th>
</tr>
</thead>
<tbody>
<tr>
<td>AES-like, Fides</td>
<td>1.125</td>
<td>1.57</td>
</tr>
<tr>
<td>0</td>
<td>&gt;50</td>
<td>80</td>
</tr>
</tbody>
</table>

near-affine relationship
Other Ciphers?

All ciphers are the same!

MI = a PRECISE and RELIABLE measure of quality of ciphers!

*note: the more repetition, like 8 or 10, the more we approach the concept of “space trails”

Example:
Compare 3 versions of DES on MI and \#SRG_2 mappings.

<table>
<thead>
<tr>
<th>U1</th>
<th>U2</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>13</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>total</td>
<td>255</td>
<td>total</td>
<td>505</td>
<td>total</td>
<td>236</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# of SRG_2 forbidden mappings
Related Work

Tezcan 2014:

For a specific input difference of an S-box, some bits of the output difference remain invariant...

<table>
<thead>
<tr>
<th>Input Difference</th>
<th>Output Difference</th>
<th>Input Difference</th>
<th>Output Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>00001</td>
<td>01?0?</td>
<td>10000</td>
<td>010?0</td>
</tr>
<tr>
<td>00010</td>
<td>1???1</td>
<td>10001</td>
<td>10??1</td>
</tr>
<tr>
<td>00011</td>
<td>0??0?</td>
<td>10011</td>
<td>0??0</td>
</tr>
<tr>
<td>00100</td>
<td>01?10</td>
<td>10100</td>
<td>01?1</td>
</tr>
<tr>
<td>00101</td>
<td>0???1</td>
<td>10101</td>
<td>0???1</td>
</tr>
<tr>
<td>00110</td>
<td>1???1</td>
<td>10110</td>
<td>1???1</td>
</tr>
<tr>
<td>00111</td>
<td>0???1</td>
<td>10111</td>
<td>0???0</td>
</tr>
<tr>
<td>01000</td>
<td>01?1?</td>
<td>11000</td>
<td>????1</td>
</tr>
<tr>
<td>01011</td>
<td>0???1</td>
<td>11100</td>
<td>?0?0</td>
</tr>
<tr>
<td>01100</td>
<td>00?0?</td>
<td>11110</td>
<td>01???</td>
</tr>
<tr>
<td>01110</td>
<td>0?0??</td>
<td>11111</td>
<td>0???</td>
</tr>
<tr>
<td>01111</td>
<td>0?0??</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
New! Combinatorial Explosion of Undesirable Properties - 2 S-boxes

Same SRG<sub>2</sub> mappings. 2 active boxes.
Output limited to
HW=6 active bits / 30 max.

=>the same type of prediction law is expected to hold for 3,4,5,... S-boxes and for most of the attacks we study: they are all based on the same basic events.
Random Permutations, APN and Ascon S-box

A new way of classifying S-boxes from strong to weak on a 2D scale.

CLAIM: we need to contemplate the large distance which separates the Ascon S-box (MI=1.91) and an ideal S-box not in terms of differential and linear properties (the distance seems small) but in terms of:

- How hard it is for a RP to move to this area - close to impossible!
- The combinatorial explosion of undesirable properties [previous slide].
Executive Summary:

Many cryptanalytics attacks can be MAPPED to combinations of discrete combinatorial events which are pure information theoretic events: they correspond to a certain probability mass of events inside a small finite number of small scale undesirable local linearity properties.

It is possible to see that Ascon is unnecessarily weak:

1. Many tiny S-boxes are somewhat inherently weak: like 42 => 122 bits of Mutual Information
2. Large MI => prediction capability => combinatorial explosion in # undesirable properties (faster than linear).
3. We claim that the Ascon S-box is an unfortunate choice, and it never was optimized in the full light of how it leads to undesirable properties. Needs some more work.

Open problem: is there a NL layer for Ascon which simultaneously:

- has lower HW cost and low depth (possibly avoiding any XORs which are slow).
- is easy to protect against side channels and has a reversible Toffoli implementation
- has a much lower information theoretic security measure of MI for I/O differences.
- has zero or near zero undesirable mappings.

We are willing to work on any new Ascon update/proposal and to evaluate it against enclosed concerns.