WXFデータの分析

相関係数について

初版 2008年11月8日（土）
誤記修正 2008年11月26日（水），
WEB公開用修正版 2009年2月2日（月），再修正 2009年2月3日（火），再々修正 2009年2月14日（土）

・ここで “WXFデータ” とは，サイドチャネル攻撃実験データの交換用フォーマット Waveform data eXchange Format （WXF）に従って作成されたデータ．
・WXFに関する資料は，http://ipsr.ynu.ac.jp/wxf/index.html にて公開されている．

WXFデータの分析：相関係数

2008年11月08日 高橋

目的

暗号処理中の電圧や電磁波等の変動を測定した時系列データ(以下，波形という)に含まれるサイドチャネル情報の指標値(グラフの縦軸)として，波形と中間値の“相関係数”を利用することが10月3日の電力解析実験WGにて提案され，(少なくとも)1つの処理手順やアルゴリズム等を定めて実験データを分析することになった．

そこで当面の実験対象と考えられる128bitブロック暗号AESを例に，“相関係数”を求めるMatlab関数を作成した．

AESの実装と暗号化処理時の波形

SASEBO-R搭載の暗号LSIに実装されているAESは，128bitのデータレジスタを更新することで1ラウンド分の回路が動作するループアーキテクチャになっている(図1)．

SASEBO-RのAES動作時の電圧変動を測定した波形を図2に示す．クロック毎にレジスタが更新されて，AESの1ラウンドが動作し，そのタイミングで波形に大きな変動が生じていることが分る．

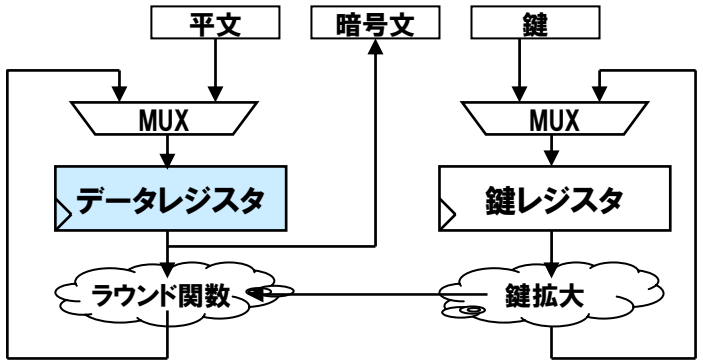


図1：AESのデータパス概略

波形と相関する‘中間値’

図2の11個目の変動を例にすると，この変動はデータレジスタが10番目の値(Drg10)から最後の値(Drg11)に更新されることで発生するものである．したがって，波形のこの部分はDrg10とDrg11とから求められる値と相関がある可能性がある．そのような値としてハミング重みH_Wとハミング距離H_Dの2つがある．

しかし波形が，どのような値と相関するのか否かは実装によって異なり，また波形と相関する中間値は，H_WとH_Dの2つだけで十分だとは言えない．

そこで，データレジスタの値をもとに中間値を求める関数に，パラメータを幾つか持たせて，今後，拡張可能な形にした．

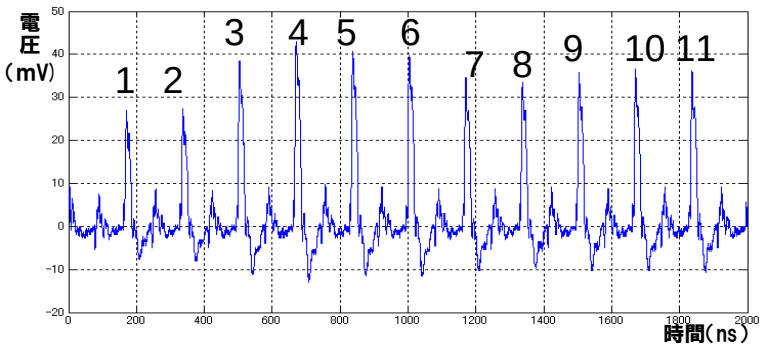


図2：AES 6MHz動作 GND側で測定

※ ある実装に対して，波形と相関係数が最大となる中間値を求める関数を特定すること(あるいはその不存在を示すこと)は，今後検討すべき課題だと考えた．

WXFデータの分析：相関係数

波形と中間値の‘相関係数’（CPAtrace）

相関係数の定義に使用する変数をここで改めて次のように定義する．

Wav　　：n個の入力（平文/暗号文）の暗号化/復号処理時に測定した t サンプルからなる波形をWav(1:t, 1:n) とする．

K　　：鍵Kを一つ固定する．
Drg　　：n個の入力Drg1 (16byte) をDrg(1:16, 1, 1:n)とする．
　　：鍵Kと入力Drg1によって決まる各ラウンドDrg2～Drg10と出力Drg11を，Drg(1:16, 2:11, 1:n)とする．
H　　：Drgの値から求めた n個の中間値を H(1:n)とする．中間値の求め方は後述する．

波形Wavと中間値Hとの“相関係数” CORを次にように求める：

COR　　：波形Wavの i サンプル目 Wav(i, 1:n)と，H(1:n)の相関係数を COR(i)とし，i = 1～t までを求める．
　　：ここで相関係数は，ピアソンの積率相関係数とする．

$$\text{ピアソンの相関係数} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

このCORは，時系列のベクトルデータである．DPAにおけるDPAtraceを意識してこのように定めた．
CORをスカラー値で代表させる場合には，次のようにする．

max(abs(COR(a:b)))　　：時間軸でa:b間の COR の絶対値の最大．
mean(COR(a:b))　　：時間軸でa:b間の COR の平均値
　　：（ただし波形Wavを平均したのちCORを求める方が高速である）．

COR(H(:), mean(Wav(a:b, :), 1))

WXFデータの分析：相関係数

中間値のパラメータ表示

AESの暗号化処理のラウンド関数の入力/出力とSASEBO-Rのデータレジスタの値の関係を図3に示す．データレジスタの初期値(平文)をDrg1，2番目の値をDrg2，10番目の値をDrg10，最後の値(暗号文)をDrg11とする．

この値から中間値として，ハミング重みH_Wとハミング距離H_Dを求める．

Drg1～Drg11により，H_WがHW1～HW11までの11個，H_DがHD2～HD11までの10個を作れる．さらに，HW，HDについて128bit全体，8bit毎，1bit毎の3段階の値を求める．8bit毎はCPAを，1bit毎はDPAを考慮した値である．128bit全体は1個，8bit毎にすると16個，1bit毎にすると128個を作れる．H_Wで1595個，H_Dで1450個，全部で3045個の中間値ができる．

この中間値を次のように表すことにする．

Xp1_p2_p3　：X は，HW または HD
：p1は，　HWのときは1～11，
：　　HDのときは2～11
：p2は，　1(128bit)，2(8bit)，3(1bit)
：p3は，　p2が1のときは1，
：　　p2が2のときは1～16，
：　　p2が3のときは1～128

先に定めた“相関係数”CORについて，使用した中間値Hを明示するためには，COR_H あるいは COR_H と表記することにする．

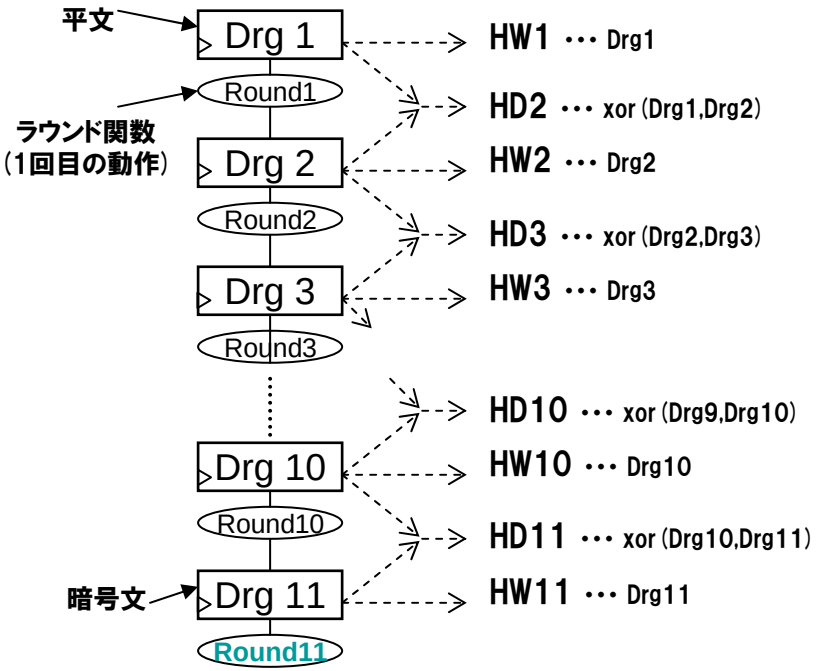
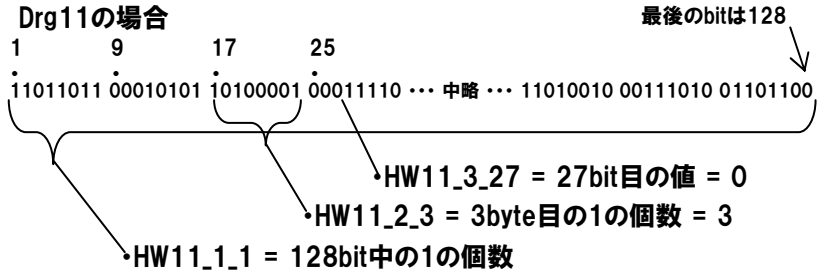


図3：AESアルゴリズムと中間値



WXFデータの分析：相関係数

処理手順の案（using Matlab）

“相関係数”を求めるMatlabのスクリプトを示す．この手順はリファレンス的なものであり，必ずしも処理時間やリソースを最適化したものではない．

図4と図5に，ハミング重みH_W，ハミング距離H_Dを求める関数を示す．図6に相関係数CORを求める関数を示す．
WXF_COR はMatlabのcorrcoef関数をそのまま使用するとメモリ不足になることがあったため用意した．

最後に図7にこれらの関数を使用して相関係数を求めるスクリプトを示す．“Wavs_A01.mat” ～ “Wavs_A05.mat” はSASEBO-RのAES実装のうち5個を測定したWXFデータから，前処理でRound11の部分を切り出したものである．サンプリングレートは5GS/sで，約6万5千個の波形がある．そのうちの1万個ほどを使うと処理時間は35秒，6万5千個全部を使用すると162秒かかる．

このスクリプトの実行結果を含め，いくらかのグラフを次ページに示す．

次ページの図8で，
AES01（AES_Comp）やAES02（AES_TBL）は，HD11のみが相関を示している．一方，AES03（AES_PPRM1）は，HW10，HW11，HD11の3本とも相関係数が 0.4～0.6 となっている．実装によって有効な中間値が異なる．

本スクリプトはMatlab用に作成したものであるが，図4～7のスクリプトはそのままでOctaveでも動作し，図8と同様なグラフが表示できる．

```
function H = WXF_HW(Drg, p1, p2, p3)
t1=reshape(dec2bin(Drg(1:16,p1,:))'-48,128,size(Drg,3));
switch p2,
case 1, H=sum(t1);
case 2, H=sum(t1(p3*8-8+[1:8],:));
case 3, H=t1(p3,:);
end;
```

ここに追加（,8）

図4：HWを求める関数

```
function H = WXF_HD(Drg, p1, p2, p3)
t1=reshape(dec2bin(Drg(1:16,p1-1,:))'-48,128,size(Drg,3));
t2=reshape(dec2bin(Drg(1:16,p1,:))'-48,128,size(Drg,3));
switch p2,
case 1, H=sum(xor(t1,t2));
case 2, H=sum(xor(t1(p3*8-8+[1:8],:),t2(p3*8-8+[1:8],:)));
case 3, H=xor(t1(p3,:),t2(p3,:));
end;
```

ここに追加（,8）

図5：HDを求める関数

```
function C=WXF_COR_v2(H, Wav)
n=size(Wav,1); C=zeros(n,1); hc=H-mean(H); c11=hc'*hc;
for t1=1:n,
    wc=Wav(t1,:)'-mean(Wav(t1,:)); c12=hc'*wc; c22=wc'*wc;
    C(t1)=c12/(sqrt(c11)*sqrt(c22));
end;
```

図6：CORを求める関数（Ver2）

```
wr=400; ws=1024*10; tar_rnd=11; cor_typ=1; cor_num=1;
load('Drgs.mat'); Drg=Drg(1:16,1:11,1:ws);
for aes=1:5; disp(['AES0',num2str(aes)]);
  load(['Wavs_A0',num2str(aes),'.mat']); Wav=Wav(1:wr,1:ws);
  H(:,1)=WXF_HW(Drg,tar_rnd-1,cor_typ,cor_num);
  H(:,2)=WXF_HW(Drg,tar_rnd  ,cor_typ,cor_num);
  H(:,3)=WXF_HD(Drg,tar_rnd  ,cor_typ,cor_num);
  for i=1:3, C(:,i)=WXF_COR(H(1:ws,i),Wav(:,1:ws)); end;
  subplot(1,5,aes); plot([1:wr]/5,C,'-');
  title(['AES0',num2str(aes),': H',num2str(tar_rnd),'-', ...
    num2str(cor_typ),'-',num2str(cor_num)]);
  axis([0 wr/5 -1 1]); grid('on'); pause(1);
end;
```

図7：WXF_HW,HD,CORの使用例

WXFデータの分析：相関係数

2008年11月08日 高橋

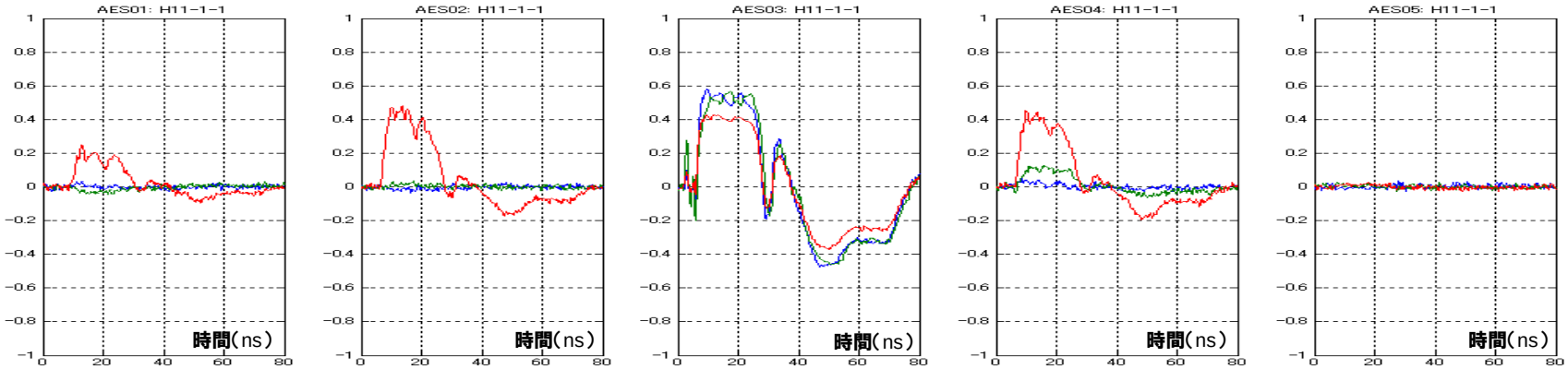


図8：図7の実行結果（AES01～AES05の波形と，128bit全体のHW，HDとの相関係数）
AES01=Comp，02=TBL，03=PPRM1，04=PPRM3，05=RSL

青:HW10，緑:HW11，赤:HD11

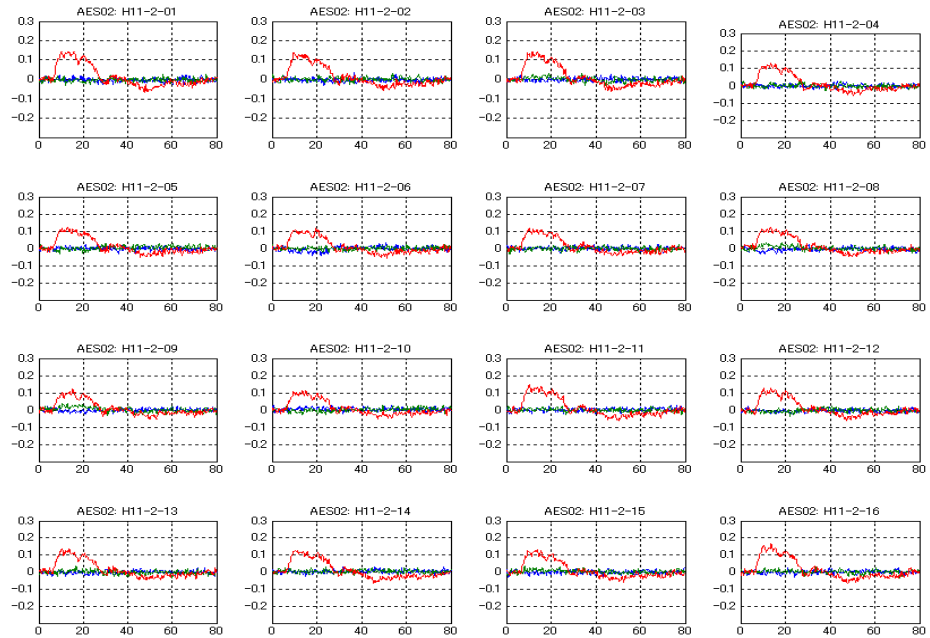


図9：AES02の波形と，8bit毎のHW，HDとの相関係数のグラフ

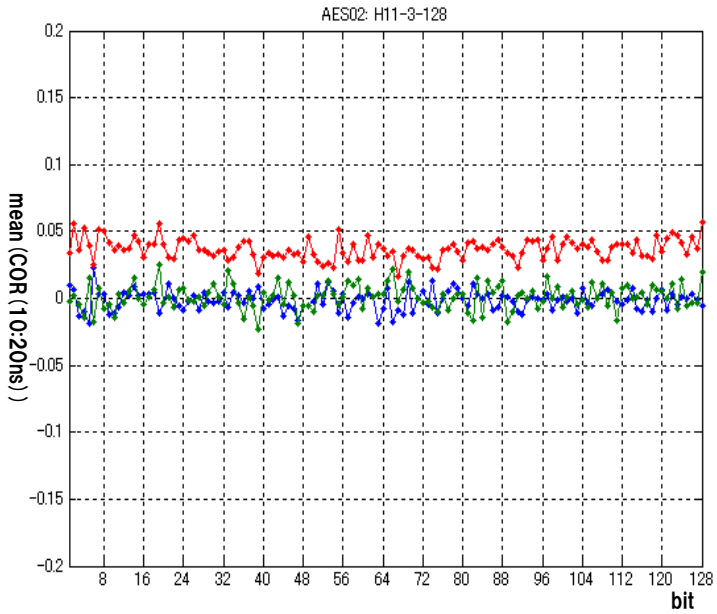


図10：AES02の波形と，1bit毎のHW，HDとの相関係数
縦軸は10～20nsの区間の平均値，横軸はbit位置．

WXFデータの分析：相関係数

波形数を変化させたときの相関係数の様子を図11に示す．波形はAES03のRound11の区間で，中間値はHW10,HW11,HD11とHD9の4つで，CORの10～20nsに対応する部分の平均値を縦軸にプロットしてある．HW10,HW11,HD11の3つは，0.4～0.6付近に収束し，HD9はRound11の波形とは相関がないので0に収束している．

HD11とHD9が分離する波形数は，CPAで正解鍵が求まる波形数に関係すると考えられる．

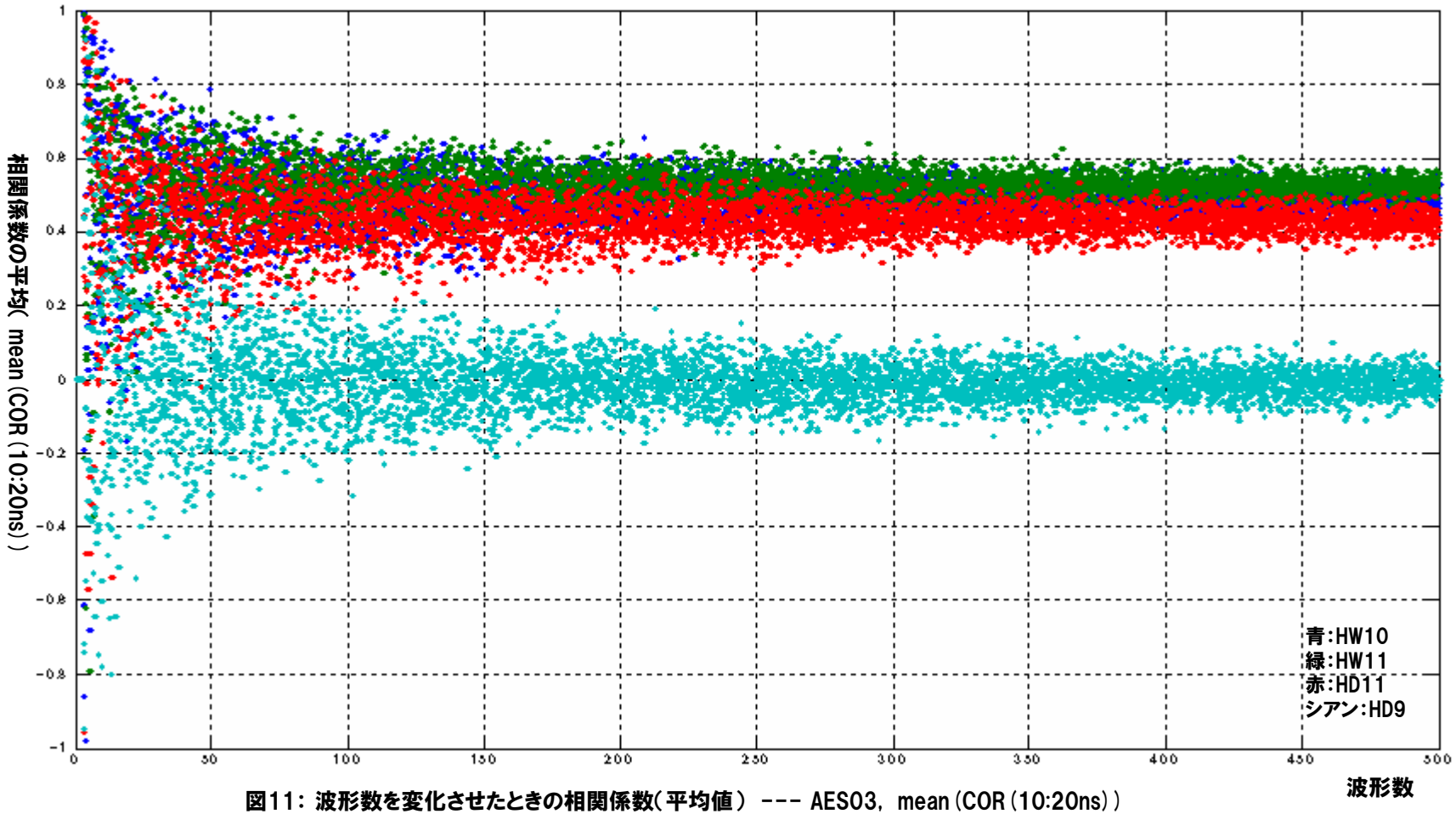


図11：波形数を変化させたときの相関係数(平均値） --- AES03, mean(COR(10:20ns))

WXFデータの分析：相関係数

2009年02月14日 高橋

まとめと今後の課題

波形と中間値との“相関係数”をサイドチャネル情報の指標値（グラフの縦軸）に利用するため，その求め方について検討した．

1．波形 Wav(1:t,1:n) と中間値 H(1:n) から，時系列ベクトルとして，相関係数 COR(1:t) を定めた．

2．CORをスカラー値で代表させる場合には，次の3タイプの変換方法がある．

（1）max(abs(COR(a:b)))　　：時間軸でa:b間の COR の絶対値の最大．
（2）mean(COR(a:b))　　：時間軸でa:b間の COR の平均値．
（3）COR(H(:), mean(Wav(a:b, :), 1))　：Hと，a:b間の波形の平均との相関係数．

オリジナルのDPAアルゴリズムでは(1)が採用されているが，(2)の方がエラーレートが低く，(3)の方が処理時間が少ない．
ただし(2)や(3)はa:bを適切に定める必要がある．

3．中間値 H として具体的に，あるAES実装を想定して，ハミング重みとハミング距離（あるいは“値型”と“遷移型”）の2つについて，各々1bit，8bit，128bit単位の3レベルの値を，AESのデータレジスタ Drg から求める Matlabスクリプトを示した．
そして，AESの5種類の実装を測定して，CORのグラフを幾つか例示した．

中間値 H は，暗号文と正しい秘密鍵があれば求めることができ，正しい鍵で求めたHは，Wavの適切な区間で高い相関を示し，間違った鍵の場合には相関が低くなるように決める．そして鍵を総当りにしてCORが最大となる値を探すことで，鍵を推測する．この最尤推定による秘密パラメータの導出がDPAやCPAの動作原理である．
※ 犯罪利用を避けるために，この鍵を総当りでHを求める関数SF(sk, Drg11)のコードは公開しない．

4．CORを用いた指標値の1案として，正しい鍵で求めたCORと，間違った鍵で求めたCORのクロスポイントを提案した．

AESの場合には8bit単位で鍵推定できるので，1本の正しい鍵によるCORと255本の間違った鍵によるCORの分布が一定の確率（95%あるいは99%など）で分離する最小波形数を，DPA型のサイドチャネル攻撃の指標値にする提案である．
“一定の確率“ は（従来のグラフの縦軸である）エラーレートに対応する値と考えられる（適当な数式で変換可能？）．

今後の課題として，
A) DPAアルゴリズムの収集．例えば標準波形データの整備とそれを使ったDPAコンテストの企画実施．
B) Aで集めたアルゴリズムと測定環境の関係の分析．測定環境と分析手段の組合せを網羅的に調査．
C) A,Bから，少なくとも既存の攻撃法と測定手段を組み合わせた範囲では安全であると言えるセキュリティ要件を定める．
などが考えられる．

WXFデータの分析：相関係数

補足

WXFデータからCORのグラフまでの流れを図12に示す．

図中のWXFデータ（AES_Comp～AES_SSS1までの5個）と
そこから生成した
Wavs.mat（Wavs_A01.mat～Wavs_A05.matの5個）と
Drgs.mat を公開します．下記ページを参照して下さい．

URL http://ipsr.ynu.ac.jp/wxf/index.html

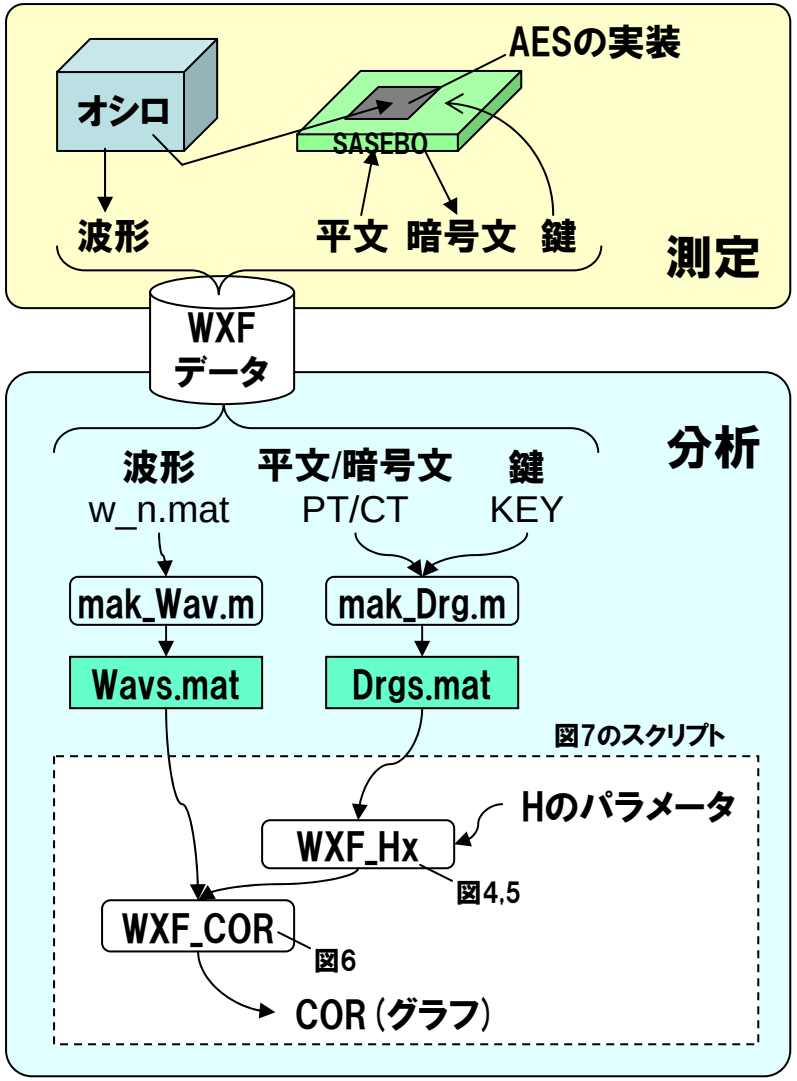


図12：WXFデータからCORグラフまでの関係

参考文献

[SSBR] 産業技術総合研究所 情報セキュリティ研究センタ，“サイドチャネル攻撃用標準評価ボードSASEBO-R”，URL http://www.rcis.aist.go.jp/special/SASEBO/SASEBO-R-ja.html

[K08] 防衛大学校 コンピュータ学研究室，SASEBO-Rの波形数と正解ブロック数のグラフ，第2回 電力解析実験WG 配布資料，資料6， 2008年10月．

[SHAS08] 菅原 健，本間 尚文，青木 孝文， 佐藤 証，“標準評価基板上のASICへの差分電力解析実験”， コンピュータセキュリティシンポジウム2008，論文集［第一分冊］，pp.533-538，D5-3，2008年10月．

[WXF08] 横浜国大 松本研究室，“交換用標準フォーマット（WXF）”，
URL http://ipsr.ynu.ac.jp/wxf/index.html

本ファイルはCRYPTREC暗号モジュール委員会の第3回電力解析実験ワーキンググループ（11月26日開催）での説明資料を，WEB公開用に編集（誤記訂正及び削除・加筆）したものである．