Methods for Symmetric Key Cryptography and Cryptanalysis

EWM PhD Summer School, Turku, June 2009

Kaisa Nyberg

kaisa.nyberg@tkk.fi

Department of Information and Computer Science

Helsinki University of Technology

and Nokia, Finland

This lecture is dedicated to the memory of

Professor Susanne Dierolf

a dear and supporting friend, a highly respected colleague,

and a great European Woman in Mathematics,

who passed away in May 2009 at the age of 64

in Trier, Germany.

Outline

1. Boolean function
   - Linear approximation of Boolean function
   - Related probability distribution
2. Cryptographic encryption primitives
   - Linear approximation of block cipher
   - Linear approximation of stream cipher
3. Cryptanalysis and attack scenarios
   - Key information deduction on block cipher
   - Distinguishing attack on stream cipher
   - Initial state recovery of stream cipher
4. Conclusions

Boolean Functions

Binary vector space

- $\mathbb{Z}_2^n$ the space of n-dimensional binary vectors
- $\oplus$ sum modulo 2
- Given two vectors

$$a = (a^1, \dots, a^n),\ b = (b^1, \dots, b^n) \in \mathbb{Z}_2^n$$

the inner product (dot product) is defined as

$$a \cdot b = a^1 b^1 \oplus \cdots \oplus a^n b^n.$$

- Then a is called the linear mask of b.

Boolean function

- $f : \mathbb{Z}_2^n \mapsto \mathbb{Z}_2$ Boolean function.
- Linear Boolean function is of the form $f(x) = u \cdot x$ for some fixed linear mask $u \in \mathbb{Z}_2^n$.
- $f : \mathbb{Z}_2^n \mapsto \mathbb{Z}_2^m$ with $f = (f_1, \ldots, f_m)$, where f_i are Boolean functions, is called a vector Boolean function of dimension m.
- A linear vector Boolean function from $\mathbb{Z}_2^n$ to $\mathbb{Z}_2^m$ is represented by an $m \times n$ binary matrix U. The m rows of U are denoted by $u_1, \ldots, u_m$, where each u_i is a linear mask.

Correlation

- The correlation between two Boolean functions $f : \mathbb{Z}_2^n \mapsto \mathbb{Z}_2$ and $g : \mathbb{Z}_2^n \mapsto \mathbb{Z}_2$ is defined as

$$c(f,g) = 2^{-n}\left(\#\{x \in \mathbb{Z}_2^n \mid f(x) = g(x)\} - \#\{x \in \mathbb{Z}_2^n \mid f(x) \neq g(x)\}\right)$$

- Correlation $c(f,0)$ is called the correlation (sometimes aka bias) of f.
- Linear cryptanalysis makes use of large correlations of Boolean functions in cipher constructions.

Random variable related to Boolean function

- X discrete random variable taking on values in $\mathbb{Z}_2^n$
- If $p = (p_\eta)_{\eta \in \mathbb{Z}_2^n}$ is the probability distribution (p.d.) of X, where $p_\eta = \Pr(X = \eta)$, we denote $X \sim p$.
- Let θ denote the uniform distribution on $\mathbb{Z}_2^n$.
- Let $f : \mathbb{Z}_2^n \mapsto \mathbb{Z}_2^m$ be a Boolean function and $X \sim \theta$. Then the p.d. of $f(X)$ is called the p.d. of f.

Walsh-Hadamard Transform

Walsh-Hadamard transform is a type of discrete Fourier-transform.

Let $\phi : \mathbb{Z}_2^n \to \mathbb{R}$ be a real-valued function. The Walsh-Hadamard transform $\widehat{\phi}$ of ϕ is defined as

$$\widehat{\phi}(u) = \sum_{x \in \mathbb{Z}_2^n} \phi(x)(-1)^{x \cdot u}, u \in \mathbb{Z}_2^n.$$

Then

$$\phi(x) = 2^{-n}\widehat{\widehat{\phi}}(x), x \in \mathbb{Z}_2^n,$$

using the inverse of Walsh-Hadamard transform.

Convolution

The convolution of two functions $\phi : \mathbb{Z}_2^n \to \mathbb{R}$ and $\psi : \mathbb{Z}_2^n \to \mathbb{R}$ is defined as

$$(\phi * \psi)(y) = \sum_{x \in \mathbb{Z}_2^n} \phi(x)\psi(x+y), y \in \mathbb{Z}_2^n.$$

Then

$$\widehat{(\phi * \psi)}(u) = \widehat{\phi}(u)\widehat{\psi}(u), u \in \mathbb{Z}_2^n.$$

Correlation and probability distribution

The correlations of masked Boolean function can be computed as Walsh-Hadamard transform of the distribution of the function:

$$c(a \cdot f) = 2^{-n} \sum_{x \in \mathbb{Z}_2^n} (-1)^{a \cdot f(x)} = \sum_{\eta \in \mathbb{Z}_2^m} (-1)^{a \cdot \eta} p_\eta = \widehat{p}(a).$$

Theorem 1 *Let $f : \mathbb{Z}_2^n \mapsto \mathbb{Z}_2^m$ be a Boolean function with p.d. p and one-dimensional correlations $c(a \cdot f)$, $a \in \mathbb{Z}_2^m$. Then*

$$p_\eta = 2^{-m} \sum_{a \in \mathbb{Z}_2^m} (-1)^{a \cdot \eta} c(a \cdot f)$$

for all $\eta \in \mathbb{Z}_2^m$.

Cryptographic Encryption Primitives

Symmetric-key encryption

$K \in \mathcal{K}$	the key
$x \in \mathcal{P}$	the plaintext
$y \in \mathcal{C}$	the ciphertext

Encryption method is a family $\{E_K\}$ of transformations $E_K : \mathcal{P} \to \mathcal{C}$, parametrised using the key K such that for each encryption transformation E_K there is a decryption transformation $D_K : \mathcal{C} \to \mathcal{P}$, such that $D_K(E_K(x)) = x$, for all $x \in \mathcal{P}$.

Block cipher

The data to be encrypted is split into blocks x_i, $i = 1, \ldots, N$ of fixed length n. A typical value of n is 128.

$\mathcal{P} = \mathcal{C} = \mathbb{Z}_2^n$, $\mathcal{K} = \mathbb{Z}_2^\ell$

Block cipher seen as a vector Boolean function

$$f : \mathbb{Z}_2^n \times \mathbb{Z}_2^\ell \to \mathbb{Z}_2^n \times \mathbb{Z}_2^n \times \mathbb{Z}_2^\ell$$

$$f(x, K) = (x, E_K(x), K)$$

Linear approximation of a block cipher

$$u \cdot x \oplus w \cdot E_K(x) \oplus v \cdot K$$

where $x \sim \theta$ and K is fixed.

Stream cipher

Data to be encrypted is split into blocks

$x_i,\ i = 1, \ldots, N$

of fixed length n. Now typical values of n are $n = 1, 8$, or 32.

$\mathcal{K} = \mathbb{Z}_2^{\ell}$

The key $K \in \mathcal{K}$ determines the initial state of a keystream generator which produces a new fresh key K_i, $i = 1, \ldots, N$, for each data block.

$\mathcal{P} = \mathcal{C} = (\mathbb{Z}_2^n)^N$

where N can be any positive integer less than the period of the keystream generator.

$E_K(x_1, \ldots, x_N) = K_1 \oplus x_1, \ldots, K_N \oplus x_N$

Linear approximation of stream cipher

Key stream generator seen as Boolean functions

$$f_i : \mathbb{Z}_2^\ell \to \mathbb{Z}_2^\ell \times \mathbb{Z}_2^n, \; f_i(K) = (K, K_i)$$

Linear approximations of key stream generator

$$u_i \cdot K \oplus w \cdot K_i$$

Cryptanalysis

Attack scenarios

Assumptions about data available to attacker

- Ciphertext only (Shannon's model)
- Known plaintext-ciphertext pairs
- Chosen (by attacker) plaintext and corresponding ciphertext

Breaks

This classification is hierarchial. An attack is successful if its complexity is less than the complexity of exhaustive key search.

- Total break: attacker gets the key
- Instance deduction: attacker gets a clone of D_K
- Key information deduction: attacker gets partial information about the key
- Distinguishing: attacker can distinguish the cipher from a purely random function

Distinguishing leads sometimes to information deduction.

Linear cryptanalysis

- Linear cryptanalysis is a known plaintext-ciphertext attack
- Linear cryptanalysis makes use of linear approximations of the cipher.
- Linear cryptanalysis can be used in distinguishing attacks or in key information deduction.

Key information recovery on block cipher

Linear approximation of a block cipher

$$u \cdot x \oplus w \cdot E_K(x) \oplus v \cdot K$$

where $x \sim \theta$ and K is fixed.

The correlation c of this Boolean function is assumed to be known or a sufficiently accurate estimate is available.

Observe a number N of known plaintext-ciphertext pairs $(x, E_K(x))$ and calculate the observed correlation $\tilde{c}$ of $u \cdot x \oplus w \cdot E_K(x)$.

Determine $v \cdot K = 0$, if $c\tilde{c} > 0$, and $v \cdot K = 1$, otherwise.

The probability of success

Consider the case $c > 0$ and $v \cdot K = 0$. Other cases are similar.

Let N_0 be the observed number of plaintexts x such that $u \cdot x \oplus w \cdot E_K(x) = 0$.

Then N_0 is binomially distributed with expected value Np and variance $Np(1-p)$, where $p = \frac{c+1}{2}$. Then

$$Z = \frac{N_0 - Np}{\sqrt{Np(1-p)}} \sim \mathcal{N}(0,1)$$

where $\mathcal{N}(0,1)$ is the standard normal distribution. Then the bit $v \cdot K$ is correctly determined if the observed correlation $\tilde{c}$ is positive, which happens if and only if $N_0 > N/2$, or equivalently, $Z > -c\sqrt{N}$. Hence the probability of success can be estimated as

$$1 - \Phi(-c\sqrt{N})$$

where Φ is the cumulative density function of $\mathcal{N}(0,1)$. The probability is 0.921 for $N = 1/c^2$. This gives an estimate of the number N of plaintext-ciphertext pairs for successful cryptanalysis.

Linear attacks on stream cipher

Let $f_i : K \mapsto K_i$ be of the form $g \circ f^i$, where f is a (linear) state transition function and g is a nonlinear state output function, aka filter function.

Assume that we have a strong linear approximation of g, that is

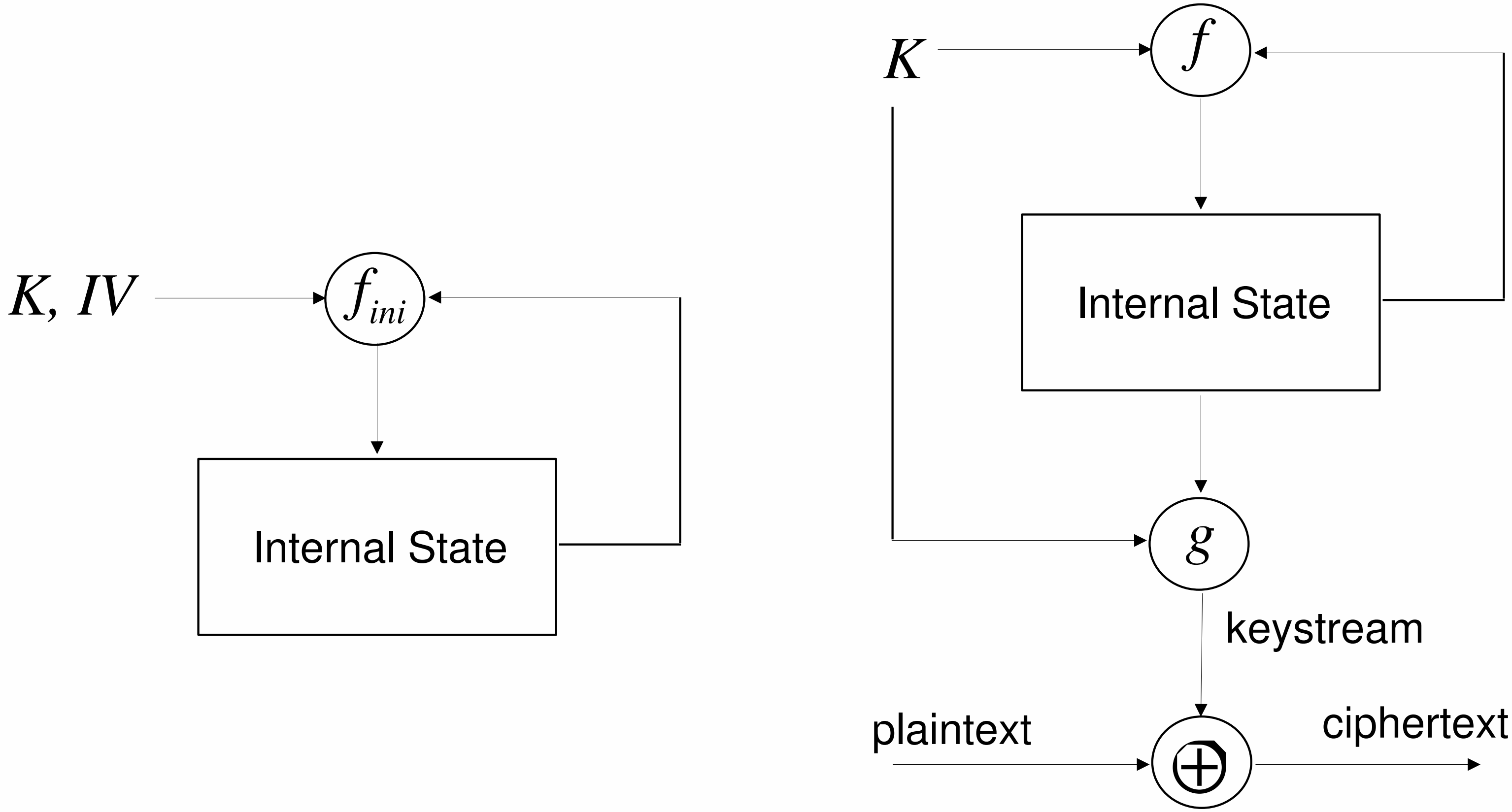
$$u \cdot x \bigoplus w \cdot g(x)$$

with correlation c.

Now two types of linear attacks can be launched:

- distinguishing attacks
- initial state information deduction attacks.

Additive synchronous stream cipher



Intialisation

Key Stream Generation and Encryption

Distinguishing attack on stream cipher

From the linear approximation we get

$$u \cdot s_i \oplus w \cdot g(s_i) = u \cdot s_i \oplus w \cdot K_i$$

with correlation c, where $s_i = f^i(K)$ is the state at time i, $i = 1, \ldots, N$.
Typically, f is a state transition function of a linear feedback shift register. Then there exist a small number of integers $a_1, \ldots a_d$ such that

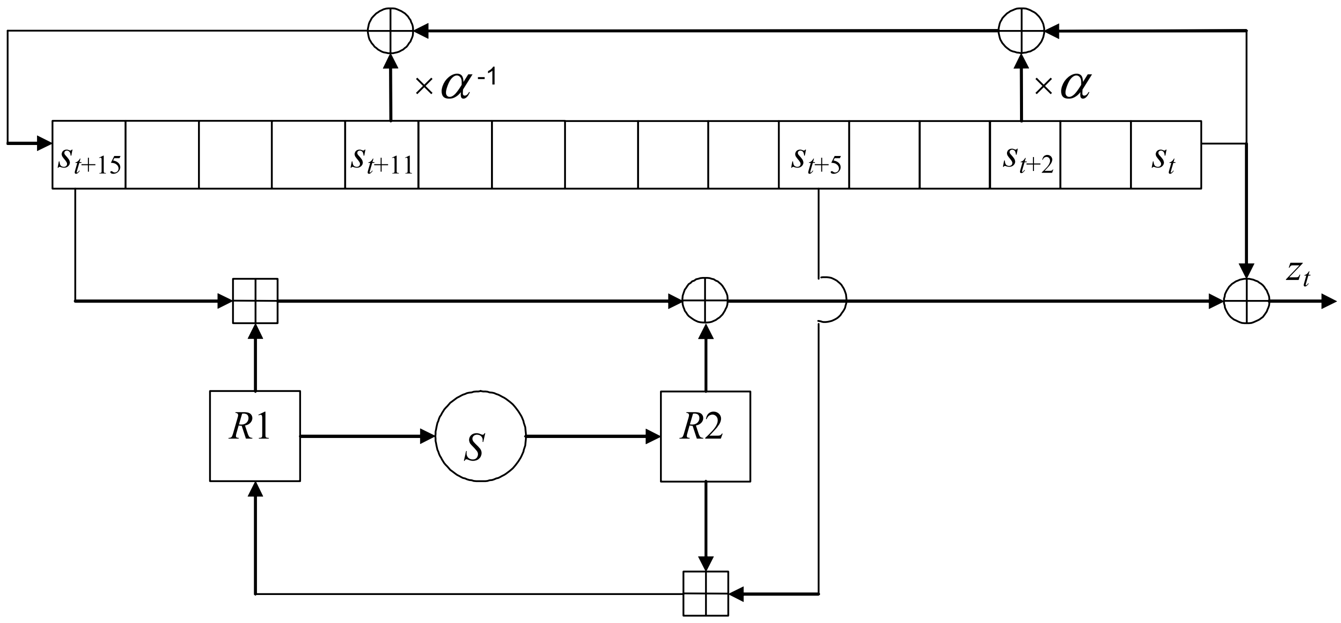
$$f^i \oplus f^{i+a_1} \oplus \ldots \oplus f^{i+a_d} = 0.$$

Then we use the linear approximation $d+1$ times to make the internal states cancel and to get a linear relation on the keystream

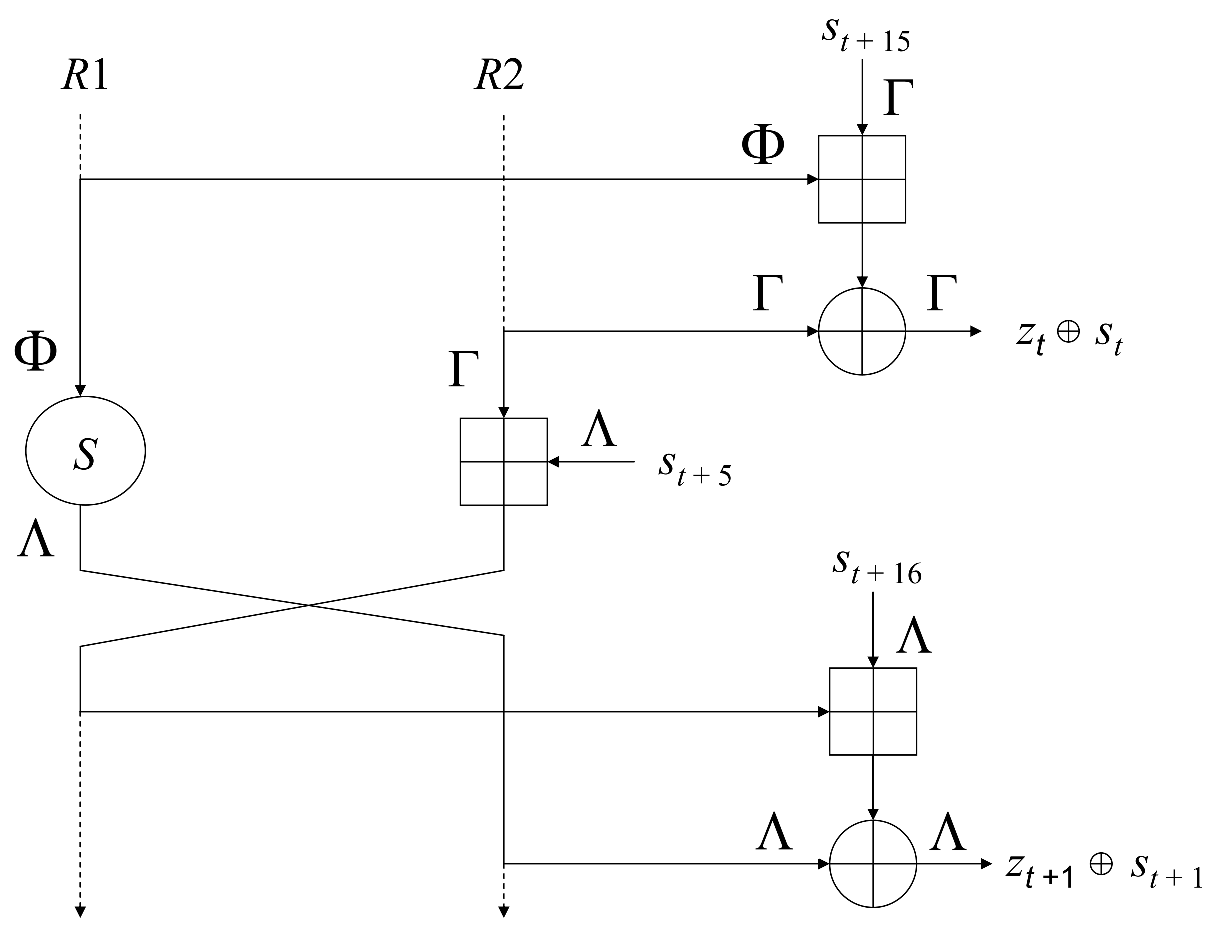
$$w \cdot (K_i \oplus K_{i+a_1} \oplus \ldots \oplus K_{i+a_d})$$

which has correlation c^{d+1}.

Snow 2.0 stream cipher



Linear approximations over Snow 2.0



Initial state recovery on stream cipher

Assume linear approximations $u \cdot s_i \oplus w \cdot g(s_i) = u \cdot s_i \oplus w \cdot K_i$ with correlation c, where $s_i = f^i(K)$ is the state at time i.

Typically, f is a state transition function of a linear feedback shift register. Let A be the transpose of f. Then we have

$$A^i u \cdot K \oplus w \cdot K_i$$

with correlation c, for all $i = 1, 2, ..., N$. Denote $b_i = w \cdot K_i$. Then the problem is to solve K from a large system of highly erroneous (but not completely random) equations

$$A^i u \cdot K = 0$$

with correlations $(-1)^{b_i} c$, for all $i = 1, 2, \ldots, N$.

A decoding problem

Given such a system

$$A^i u \cdot K = 0,$$

with correlations $(-1)^{b_i} c$, for $i = 1, 2, \ldots, N$, we can proceed as follows. Assume $c > 0$. We select $K = \eta$ such that η maximizes

$$p_\eta = \sum_{i=1}^{N} (-1)^{b_i \oplus A^i u \cdot \eta}$$

These values can be computed simultaneously for all $\eta \in \mathbb{Z}_2^\ell$ using the fast Fourier (Walsh-Hadamard) transform. The computational complexity is $\ell 2^\ell$. Indeed, no savings have been gained compared to exhaustive search of the initial state. Some savings can be achieved using a trade off between exhaustive search and the Fourier transform method.

Multidimensional linear cryptanalysis

- Makes use of a number of linear approximations simultaneously

$$Ux \oplus WE_K(x) \oplus VK$$

- Particulary useful in key information deduction attack on block ciphers: now all key information bits VK can be deduced simultaneously with (about) the same amount of data
- Can also be applied to stream cipher attacks
- The binomial statistics of one-dimensional analysis has multiple generalizations to the multidimensional case: χ^2, Log-likelihood ratio, etc.

Conclusions

- Linear cryptanalysis is one of the most powerful cryptanalytic methods.
- The best known attacks on many contemporary good ciphers are linear attacks.
- Resistance against linear cryptanalysis is one of the main design criteria for symmetric key ciphers.
- Extensions to multidimensional linear approximations have been found to bring significant enhancements.
- Decoding algorithms and techniques may be helpful in improving the efficiency of key information deduction attacks.

Literature

Mitsuru Matsui: Linear Cryptanalysis Method for DES Cipher. In Helleseth, T., ed.: Advances in Cryptology – EUROCRYPT '93. Volume 765 of Lecture Notes in Computer Science, Berlin/Heidelberg, Springer (1994) 386–397

Martin Hell, Thomas Johansson, Lennart Brynielsson: An overview of distinguishing attacks on stream ciphers. Volume 1 of Cryptography and Communications, Springer (2009)

Come Berbain, Henri Gilbert, Alexander Maximov: Cryptanalysis of Grain. In Robshaw, M., ed.: Fast Software Encryption. Volume 4047 of Lecture Notes in Computer Science, Berlin/Heidelberg, Springer (2006) 15–29

Miia Hermelin, Joo Yeon Cho, Kaisa Nyberg: Multidimensional extension of Matsui's Algorithm 2. In: Fast Software Encryption. Volume 5665 of Lecture Notes in Computer Science, Springer (2009) 209–227