INTERNATIONAL STANDARD

ISO 8731-2

First edition
1987-12-15

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ

Banking — Approved algorithm for message authentication —

Part 2 : Message authenticator algorithms

Banque — Algorithmes approuvés pour l'authentification des messages —

Partie 2 : Algorithme d'authentification des messages

Reference number
ISO 8731-2 : 1987 (E)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work.

Draft International Standards adopted by the technical committees are circulated to the member bodies for approval before their acceptance as International Standards by the ISO Council. They are approved in accordance with ISO procedures requiring at least 75 % approval by the member bodies voting.

International Standard ISO 8731-2 was prepared by Technical Committee ISO/TC 68, *Banking and related financial services.*

Users should note that all International Standards undergo revision from time to time and that any reference made herein to any other International Standard implies its latest edition, unless otherwise stated.

Printed in Switzerland

Contents

Page

1 Scope and field of application 1

2 Reference 1

3 Brief description 1

3.1 General 1

3.2 Technical 1

4 The algorithm 1

4.1 Definition of the functions used in the algorithm 1

4.2 Specification of the algorithm 3

5 Specification of the mode of operation 3

Annex

Test examples for implementation of the algorithm 5

Banking — Approved algorithm for message authentication —

Part 2 : Message authenticator algorithms

1 Scope and field of application

ISO 8731 specifies, in individual parts, approved authentication algorithms i.e. approved as meeting the authentication requirements specified in ISO 8730. This part of ISO 8731 deals with the Message Authenticator Algorithm for use in the calculation of the Message Authentication Code (MAC).

The Message Authenticator Algorithm (MAA) is specifically designed for high speed authentication using a mainframe computer. This is a special purpose algorithm to be used where data volumes are high, and efficient implementation by software a desirable characteristic. MAA is also suitable for use with a programmable calculator.

Test examples are given in an annex, which does not form part of this International Standard.

2 Reference

ISO 8730, *Banking — Requirements for message authentication (wholesale).*

3 Brief description

3.1 General

The Message Authenticator Algorithm works on the principle of a Message Authentication Code (or MAC), a number sent with a message, so that a check can be made by the receiver of the message that it has not been altered since it left the sender.

3.2 Technical

All numbers manipulated in this algorithm shall be regarded as 32-bit unsigned integers, unless otherwise stated. For such a number N, $0 \leqslant N < 2^{32}$. This algorithm can be implemented conveniently and efficiently in a computer with a word length of 32 bits or more.

Messages to be authenticated may originate as a bit string of any length. They shall be input to the algorithm as a sequence of 32 bit numbers, M_1, M_2 — M_n, of which there are n, called message blocks. The detail of how to pad out the last block M_n to 32 bits is not part of the algorithm but shall be defined in any application. This algorithm shall not be used to authenticate messages with more than 1 000 000 blocks, i.e. $n \leqslant 1\,000\,000$.

The key shall comprise two 32 bit numbers J and K and thus has a size of 64 bits.

The result of the algorithm is a 32 bit authenticator value denoted Z. The calculation can be performed on messages as short as one block ($n = 1$).

The calculation has three parts

a) The prelude shall be a calculation made with the keys (J and K) alone and it shall generate six numbers X_0, Y_0, V_0, W, S and T which shall be used in the subsequent calculations. This part need not be repeated until a new key is installed.

b) The main loop is a calculation which shall be repeated for each message block M_i and therefore, for long messages, dominates the calculation.

c) The coda shall consist of two operations of the main loop, using as its message blocks the two numbers S and T in turn, followed by a simple calculation of Z, the authenticator.

The mode of operation (see clause 5) is an essential feature of the implementation of this algorithm.

The figure shows the data flow in schematic form.

4 The algorithm

4.1 Definition of the functions used in the algorithm

4.1.1 General definitions

A number of functions are used in the description of the algorithm. In the following, X and Y are 32 bit numbers and the result is a 32 bit number except where stated otherwise.

CYC(X) is the result of a one-bit cyclic left shift of X.

AND(X,Y) is the result of the logical AND operation carried out on each of 32 bits.

OR(X,Y) is the result of the logical OR operation carried out on each of 32 bits.

XOR(X,Y) is the result of the XOR operation (modulo 2 addition) carried out on each of 32 bits.

ADD(X,Y) is the result of adding X and Y discarding any carry from the 32nd bit, that is to say, addition modulo 2^{32}.

CAR(X,Y) is the value of the carry from the 32nd bit when X is added to Y; it has the value 0 or 1.

MUL1(X,Y), MUL2(X,Y) and MUL2A(X,Y) are three different forms of multiplication, each with a 32 bit result.

4.1.2 Definition of multiplication functions

To explain the multiplications, let the 64 bit product of X and Y be [U,L]. Here the square brackets mean that the values enclosed are concatenated, U on the left of L. Hence U is the upper (most significant) half of the product and L the lower (least significant) half.

4.1.2.1 To calculate MUL1(X,Y)

Multiply X and Y to produce [U,L]. With S and C as local variables,

S := ADD(U,L); ... (1)

C := CAR(U,L); ... (2)

MUL1(X,Y) := ADD(S,C). ... (3)

That is to say, U shall be added to L with end around carry.

Numerically the result is congruent to X*Y, the product of X and Y, modulo $(2^{32} - 1)$. It is not necessarily the smallest residue because it may equal $2^{32} - 1$.

4.1.2.2 To calculate MUL2(X,Y)

This form of multiplication shall not be used in the main loop, only in the prelude. With D, E, F, S and C as local variables,

D := ADD(U,U); ... (4)

E := CAR(U,U); ... (5)

F := ADD(D,2E); ... (6)

S := ADD(F,L); ... (7)

C := CAR(F,L); ... (8)

MUL2(X,Y) := ADD(S,2C). ... (9)

Numerically the result is congruent to X*Y, the product of X and Y, modulo $(2^{32} - 2)$. It is not necessarily the smallest residue because it may equal $2^{32} - 1$ or $2^{32} - 2$.

4.1.2.3 To calculate MUL2A(X,Y)

This is a simplified form of MUL2(X,Y) used in the main loop, which yields the correct result only when at least one of the numbers X and Y has a zero in its most significant bit.

This form of multiplication is employed for economy in processing. D, S, C are local variables.

D := ADD(U,U); ... (10)

S := ADD(D,L); ... (11)

C := CAR(D,L); ... (12)

MUL2A(X,Y) := ADD(S,2C). ... (13)

The result is congruent to X*Y modulo $(2^{32} - 2)$ under the conditions stated because, in the notation of MUL2(X,Y) above, the carry E = 0.

4.1.3 Definition of the functions BYT[X,Y] and PAT[X,Y]

A procedure is used in the prelude to condition both the keys and the results in order to prevent long strings of ones or zeros. It produces two results which are the conditioned values of X and Y and a number PAT[X,Y] which records the changes that have been made. PAT[X,Y] ⩽ 255 so it is essentially an 8 bit number.

X and Y are regarded as strings of bytes. Using the notation [X,Y...] for concatenating,

$[X,Y] = [B_0, B_1, B_2, B_3, B_4, B_5, B_6, B_7]$

Thus bytes B_0 to B_3 are derived from X and B_4 to B_7 from Y.

The procedure is best described by a procedure where each byte B_i is regarded as an integer of length 8 bits.

```
begin
P := 0;
for i := 0 to 7 do
begin
    P := 2*P;
    if B[i] = 0 then
    begin
        P := P + 1;
        B'[i] := P
    end
else if B[i] = 255 then
    begin
        P := P + 1;
        B'[i] := 255 - P
    end
else
B'[i] := B[i];
end
end;
```

NOTE — The procedure is written in the programming language PASCAL (see ISO 7185), except that the non-standard identifier B' has been used to maintain continuity with the text. The symbols B[i] and B'[i] correspond to B_i and B'_i in the text.

The results are

$$BYT[X,Y] = [B'_0, B'_1, B'_2, B'_3, B'_4, B'_5, B'_6, B'_7]$$

and

$$PAT[X,Y] = P$$

4.2 Specification of the algorithm

4.2.1 The prelude

$$[J_1,K_1] := BYT[J,K];$$

$$P := PAT[J,K];$$

$$Q := (1 + P)*(1 + P). \quad \ldots (14)$$

First, by means of a calculation using J_1, produce H_4, H_6, and H_8 from which X_0, V_0 and S are derived.

$$J1_2 := MUL1(J_1,J_1); \quad J2_2 := MUL2(J_1,J_1);$$

$$J1_4 := MUL1(J1_2,J1_2); \quad J2_4 := MUL2(J2_2,J2_2);$$

$$J1_6 := MUL1(J1_2,J1_4); \quad J2_6 := MUL2(J2_2,J2_4);$$

$$J1_8 := MUL1(J1_2,J1_6); \quad J2_8 := MUL2(J2_2,J2_6). \quad \ldots (15)$$

$$H_4 := XOR(J1_4,J2_4);$$

$$H_6 := XOR(J1_6,J2_6);$$

$$H_8 := XOR(J1_8,J2_8). \quad \ldots (16)$$

From a similar calculation using K_1, produce H_5, H_7 and H_9, from which Y_0, W and T are derived.

$$K1_2 := MUL1(K_1,K_1); \quad K2_2 := MUL2(K_1,K_1);$$

$$K1_4 := MUL1(K1_2,K1_2); \quad K2_4 := MUL2(K2_2,K2_2);$$

$$K1_5 := MUL1(K_1,K1_4); \quad K2_5 := MUL2(K_1,K2_4);$$

$$K1_7 := MUL1(K1_2,K1_5); \quad K2_7 := MUL2(K2_2,K2_5);$$

$$K1_9 := MUL1(K1_2,K1_7); \quad K2_9 := MUL2(K2_2,K2_7). \quad \ldots (17)$$

$$H' := XOR(K1_5,K2_5);$$

$$H_5 := MUL2(H',Q);$$

$$H_7 := XOR(K1_7,K2_7);$$

$$H_9 := XOR(K1_9,K2_9). \quad \ldots (19)$$

Finally, condition the results using the BYT function

$$[X_0,Y_0] := BYT[H_4,H_5];$$
$$[V_0,W] := BYT[H_6,H_7];$$
$$[S,T] := BYT[H_8,H_9]. \quad \ldots (20)$$

4.2.2 The main loop

This loop shall be performed in turn for each of the message blocks M_i. In addition to M_i, the principal values employed shall be X and Y and the main results shall be the new values of X and Y. It shall also use V and W and modify V at each performance. X, Y and V shall be initialized with the values provided by the prelude. In order to use the same keys again, the initial values of X, Y and V shall be preserved, therefore they shall be denoted X_0, Y_0 and V_0 and there shall be an initializing step $X := X_0$, $Y := Y_0$, $V := V_0$, after which the main loop shall be entered for the first time. The coda, which shall be used after all message blocks have been processed by *n* cycles of the loop, is described in 4.2.3.

NOTE — The program is shown in columns to clarify its parallel operation but it should be read in normal reading order, left to right on each line.

$$V := CYC(V);$$
$$E := XOR(V,W); \quad \ldots (21)$$
$$X := XOR(X,M_i); \quad Y := XOR(Y,M_i); \quad \ldots (22)$$
$$F := ADD(E,Y); \quad G := ADD(E,X);$$
$$F := OR(F,A); \quad G := OR(G,B);$$
$$F := AND(F,C); \quad G := AND(G,D); \quad \ldots (23)$$
$$X := MUL1(X,F); \quad Y := MUL2A(Y,G). \quad \ldots (24)$$

The numbers A, B, C, D are constants which are, in hexadecimal notation :

Constant A :	0204	0801
Constant B :	0080	4021
Constant C :	BFEF	7FDF
Constant D :	7DFE	FBFF

NOTE — Lines (21) are common to both paths. Line (22) introduces the message block M_i. Lines (23) prepare the multipliers and line (24) generates new X and Y values. Only X, Y and V are modified for use in the next cycle. F and G are local variables. Since the constant D has its most significant digit zero, $G < 2^{31}$ and this ensures that MUL2A in line (24) will give the correct result.

4.2.3 The coda

After the last message block M_n has been processed, the main loop shall be performed with message block S, then again with block T, i.e. $M_{n+1} = S$, $M_{n+2} = T$.

After this, the Message Authentication Code (MAC) shall be calculated as Z = XOR(X, Y) and the algorithm shall then be complete.

NOTE — In order to calculate further MAC values without repeating the prelude (key calculation) until the keys are changed the values X_0, Y_0, V_0, W, S and T should be retained.

5 Specification of the mode of operation

Messages longer than 1 024 bytes shall be divided into blocks of 1 024 bytes and chained as follows.

For the first block of 1 024 bytes the MAC (4 bytes) shall be formed. The MAC value shall be prefixed to (but not transmitted in) the second block and the resultant 1 028 bytes authenticated. This procedure shall continue, with the MAC of each block prefixed to the next, until the last block, which need not be of size 1 024 bytes, and the final MAC shall be used as the transmitted MAC for the whole message.

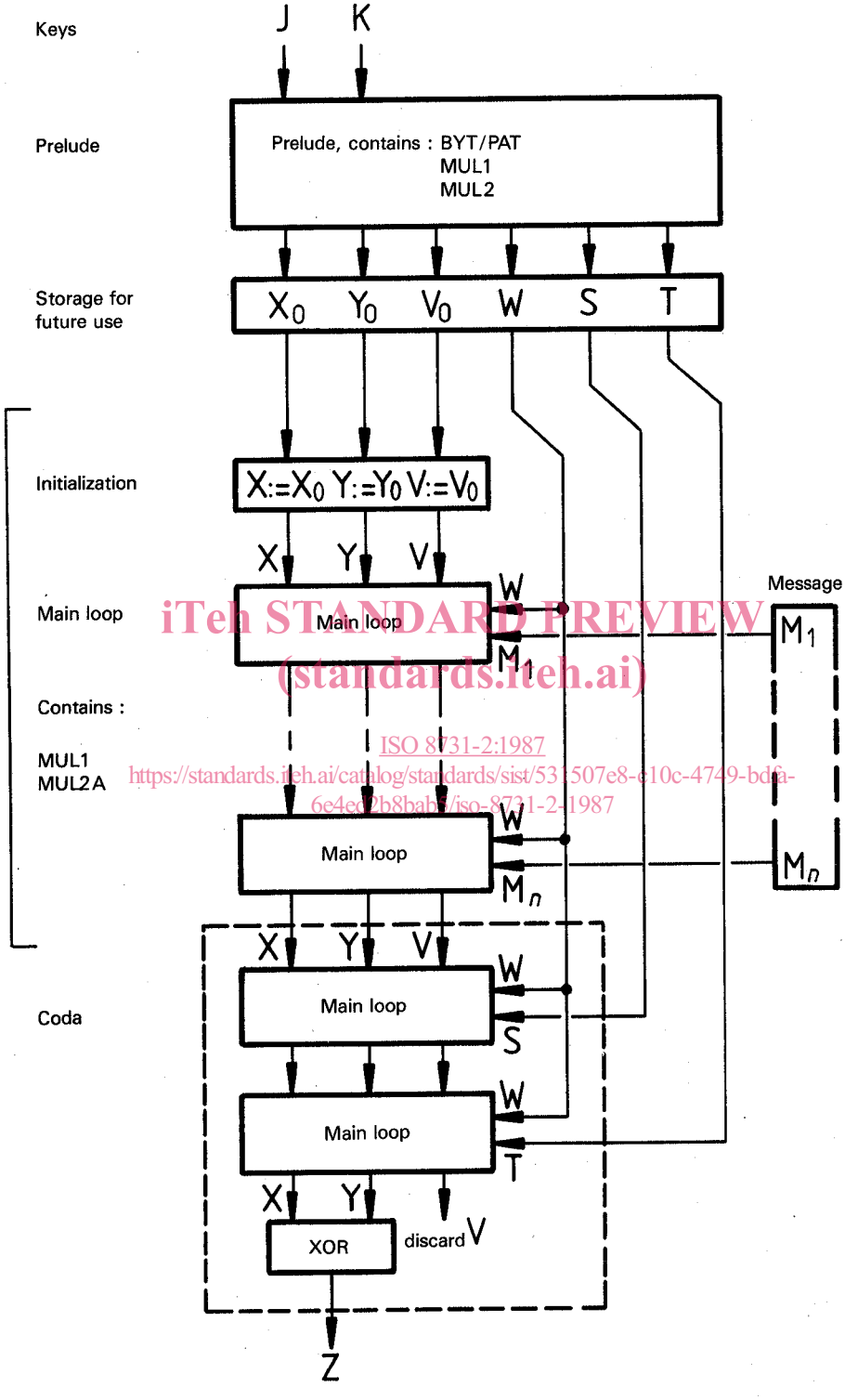

Figure — Schematic showing data flow

Annex

Test examples for implementation of the algorithm

(This annex does not form part of this Standard.)

A.0 Introduction

For most parts of the algorithm, simple test examples are given. The data used are not always realistic, i.e. they are not values which could be produced by earlier parts of the algorithm, and artificial values of constants are used. This is done to keep the test cases so simple that they can be verified by a pencil and paper calculation and thus the verification of the algorithm's implementations does not consist of comparing one machine implementation with another. The parts thus tested are

- MUL1, MUL2, MUL2A;
- BYT[X,Y] and PAT[X,Y];
- Prelude, except the initial BYT[J,K] operation;
- Main loop.

The coda is not tested separately because it uses only the main loop and one XOR function. For testing the whole algorithm, some results from a trial implementation are given.

A.1 Test examples for MUL1, MUL2, MUL2A

It is suggested that the multiplication operations should be tested with very small numbers and very large numbers. To represent a large number these examples use the ones complement. Thus if a is a small number (say less than 4 096) the notation $\overline{a}$ is used to mean its complement, i.e. $2^{32} - 1 - a$.

For small numbers a and b, all three multiplication functions produce their true product $a*b$. When large numbers are used the functions can give different results. They should be tested both ways round, with MUL(x,y) and MUL(y,x) to verify that these are equal.

A.1.1 Test cases for MUL1

In modulo $(2^{32} - 1)$ arithmetic $\overline{a}$ is effectively $-a$, therefore the results are very simple

$$\text{MUL1}(\overline{a},b) = \text{MUL1}(a,\overline{b}) = \overline{a*b}$$

$$\text{MUL1}(\overline{a},\overline{b}) = a*b$$

Examples for testing are given in table 1.

A.1.2 Test cases for MUL2

$$\text{MUL2}(\overline{a},b) = \overline{a*b - b + 1}$$

$$\text{MUL2}(a,\overline{b}) = \overline{a*b - a + 1}$$

$$\text{MUL2}(\overline{a},\overline{b}) = a*b - a - b + 1$$

Examples for testing are given in table 1.

A.1.3 Test cases for MUL2A

This will give the same result as MUL2 when tested with numbers within its range. For testing with large numbers, $\overline{a}$ and $\overline{b} - 2^{31}$ shall be used

$$\text{MUL2A}(\overline{a},b) = \overline{a*b - b + 1}$$

$$\text{MUL2A}(a,\overline{b}) = \overline{a*b - a + 1}$$

$$\text{MUL2A}(\overline{a},\overline{b} - 2^{31}) = 2^{31} * (1 - p) + a*b + p - b - 1$$

where p is the parity of a; the value of its least significant bit.