

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4680359号  
(P4680359)

(45) 発行日 平成23年5月11日 (2011.5.11)

(24) 登録日 平成23年2月10日 (2011.2.10)

(51) Int. Cl.

F I

H O 3 M 13/15 (2006.01)

H O 3 M 13/15

G O 6 F 11/10 (2006.01)

G O 6 F 11/10 3 3 O M

H O 3 M 13/09 (2006.01)

H O 3 M 13/09

H O 4 B 14/04 (2006.01)

H O 4 B 14/04

Z

H O 4 L 1/00 (2006.01)

H O 4 L 1/00

B

請求項の数 38 外国語出願 (全 50 頁)

(21) 出願番号 特願2000-282732 (P2000-282732)  
 (22) 出願日 平成12年9月18日 (2000.9.18)  
 (65) 公開番号 特開2001-189665 (P2001-189665A)  
 (43) 公開日 平成13年7月10日 (2001.7.10)  
 審査請求日 平成19年9月18日 (2007.9.18)  
 (31) 優先権主張番号 09/399,201  
 (32) 優先日 平成11年9月17日 (1999.9.17)  
 (33) 優先権主張国 米国 (US)

(73) 特許権者 501114844  
 デジタル ファウンテン、 インコーポレ  
 イテッド  
 アメリカ合衆国、カリフォルニア州 92  
 121, サン・ディエゴ、モアハウス・ド  
 ライブ 5775  
 (74) 代理人 100108855  
 弁理士 蔵田 昌俊  
 (74) 代理人 100091351  
 弁理士 河野 哲  
 (74) 代理人 100088683  
 弁理士 中村 誠  
 (74) 代理人 100109830  
 弁理士 福原 淑弘

最終頁に続く

(54) 【発明の名称】 通信システムのための情報付加的群符号発生器およびデコーダ

(57) 【特許請求の範囲】

【請求項 1】

出力記号の 1 群を生成する方法であって、1 群は 1 つ以上の出力記号であり、各出力記号は出力アルファベットから選択され、入力アルファベットからそれぞれ選択される順列の複数の入力記号を含む入力ファイルが、このような群の 1 セットから回復可能であるような該群であり、該方法は、

該群のキー I を得るステップであって、該キーがキーアルファベットから選択され、該キーアルファベットにおける選択可能なキーの数が、任意の予想される入力ファイルサイズについて、該入力ファイルにおける入力記号の数に対して、実質的に無限である、ステップと、

I を変数とする所定の関数にしたがって、該群のリスト A L ( I ) を計算するステップであって、A L ( I ) は、該群に関連付けられた入力記号に関連付けられた重み W ( I )を示し、重み W ( I ) は、少なくとも 2 つの値の間で変化する正の整数であり、少なくとも 1 つの I の値について重み W ( I ) が、1 より大きい、ステップと、

A L ( I ) によって示される該関連付けられた入力記号を変数とする所定の関数から、該群における各出力記号の出力記号値を生成するステップと、を含む方法。

【請求項 2】

前記キー I を得るステップが、ランダム関数または擬似ランダム関数にしたがってキー I を計算するステップを含む、請求項 1 に記載の方法。

## 【請求項 3】

前記計算するステップが、 $I$  のランダム関数または擬似ランダム関数にしたがって  $W(I)$  を計算するステップを含む、請求項 1 に記載の方法。

## 【請求項 4】

前記計算するステップが、 $I$  のランダム関数または擬似ランダム関数にしたがって  $AL(I)$  を計算するステップを含む、請求項 1 に記載の方法。

## 【請求項 5】

前記計算するステップが、前記群における出力記号の数  $G(I)$  を計算するステップを含み、該数  $G(I)$  が正の整数であり、該数  $G(I)$  が  $I$  のランダム関数または擬似ランダム関数にしたがって計算される、請求項 1 に記載の方法。

10

## 【請求項 6】

前記計算するステップが、  
 $I$  を変数とする所定の関数および確率分布にしたがって、重み  $W(I)$  を計算するステップであって、該確率分布が少なくとも 2 つの正の整数上にあり、そのうちの少なくとも 1 つが 1 より大きい、ステップと、  
 リスト  $AL(I)$  のリストエントリを計算するステップと、  
 $W(I)$  個のリストエントリが計算されるまで、リスト  $AL(I)$  のリストエントリを計算するステップを繰り返すステップと、  
 を含む、請求項 1 に記載の方法。

20

## 【請求項 7】

前記  $W(I)$  を計算するステップが、重み  $W(I)$  が前記キーアルファベット上の所定の分布を近似するような 重み  $W(I)$  を決定するステップを含む、請求項 6 に記載の方法。

## 【請求項 8】

前記所定の分布が一様分布である、請求項 7 に記載の方法。

## 【請求項 9】

前記所定の分布が釣り鐘曲線分布である、請求項 7 に記載の方法。

## 【請求項 10】

前記所定の分布は、 $W(I) = 1$  が  $1/K$  の確率を有し、ここで、 $K$  が前記入力ファイルにおける入力記号の数であり、そして  $W(I) = i$  が  $i = 2, \dots, K$  に対して  $1/i(i-1)$  の確率を有する、請求項 7 に記載の方法。

30

## 【請求項 11】

前記  $AL(I)$  によって示される関連付けられた入力記号を 変数とする 前記所定の関数が、リード-ソロモン符号によって決定される、請求項 1 に記載の方法。

## 【請求項 12】

前記入力アルファベットおよび前記出力アルファベットが同じアルファベットである、請求項 1 に記載の方法。

## 【請求項 13】

前記入力アルファベットが  $2^{M_i}$  個の記号を含み、各入力記号が  $M_i$  ビットを符号化し、前記出力アルファベットが  $2^{M_o}$  個の記号を含み、各出力記号の群が  $M_o$  ビットを符号化する、請求項 1 に記載の方法。

40

## 【請求項 14】

各後続キー  $I$  が 前記キーアルファベットのシーケンスにおける次のキーである、請求項 1 に記載の方法。

## 【請求項 15】

出力記号の複数の群を符号化する方法であって、出力記号の各群は請求項 1 に記載され、該方法は、

生成されるべき該出力記号の群の各々に対するキー  $I$  を生成するステップと、  
 および

該生成された出力記号の群の各々をデータレイズチャネルを介して送信されるべき出

50

カシーケンスとして出力するステップと、  
をさらに含む方法。

【請求項 16】

各キー  $I$  が他の選択されるキーとは独立に選択される、請求項 15 に記載の方法。

【請求項 17】

前記  $AL(I)$  を計算するステップが、  
前記入力ファイルにおける入力記号の数  $K$  を少なくとも近似的に識別し、かつ重み  $W(I)$  を識別するステップと、  
 $K$  より大きいまたは  $K$  に等しい最小素数  $P$  を決定するステップと、  
 $P$  が  $K$  より大きい場合、少なくとも該入力ファイルに  $P - K$  個のブランク入力記号を付加するステップと、

10

1  $X < P$  である第 1 の整数  $X$ 、および 0  $Y < P$  である第 2 の整数  $Y$  を生成するステップと、

1 ~  $W(I)$  の各  $J$  に対して、 $AL(I)$  における第  $J$  番目のエントリを  $((Y + (J - 1) \cdot X) \bmod P)$  に設定するステップと、  
を含む、請求項 1 に記載の方法。

【請求項 18】

前記各  $J$  に対して  $AL(I)$  における第  $J$  番目のエントリを設定するステップが、  
配列  $V$  における第 1 のエントリ  $V[J = 0]$  を  $Y$  に設定するステップと、  
1 ~  $W(J) - 1$  の各  $J$  に対して、該配列  $V$  における第  $J$  番目のエントリ  $V[J]$  を  $((V[J - 1] + X) \bmod P)$  に設定するステップと、  
該配列  $V$  を該リスト  $AL(I)$  として使用するステップと、  
を含む、請求項 17 に記載の方法。

20

【請求項 19】

ソースから目的地にパケット通信チャネルを介してデータを送信する方法であって、該方法が、

a) 送信されるべき該データを入力記号の順列セットとして構成するステップであって、各入力記号は、入力アルファベットから選択され、該データにおける位置を有する、ステップと、

b) 出力記号の複数の群を生成するステップであって、各出力記号が出力アルファベットから選択され、該複数の群の各群が、以下のステップによって生成され、該ステップは、

30

キーアルファベットから、生成される該群のキー  $I$  を選択するステップと、  
 $I$  の関数として重み  $W(I)$  を決定するステップであって、ここで、重み  $W(I)$  は、少なくとも 2 つの値の間および該キーアルファベット上で変化する正の整数であり、該キーアルファベットにおける少なくともいくつかのキーに対して 1 より大きい、ステップと、

$I$  の関数にしたがって  $W(I)$  個の該入力記号を選択し、したがって、該群に関連付けられる  $W(I)$  個の入力記号のリスト  $AL(I)$  を形成するステップと、

該群における出力記号の数  $G(I)$  を決定するステップと、  
該関連付けられる  $W(I)$  個の入力記号を変数とする所定の値関数から該群における各出力記号の値  $B(I)$  を計算するステップとを含む、ステップと、

40

c) 該複数の群のうちの少なくとも 1 つの出力記号の群のうちの少なくとも 1 つを、複数のパケットの各々にパケット化するステップと、

d) 該複数のパケットを、該パケット通信チャネルを介して送信するステップと、

e) 該複数のパケットのうちの少なくともいくつかを該目的地で受信するステップと、

f) 該複数の受信されたパケットから該データを復号化するステップと、  
を含む方法。

【請求項 20】

50

前記データを復号化するステップが、

- 1) 受信された出力記号の各群を処理するステップであって、
  - a) 該群の前記キー I を決定するステップと、
  - b) 該群の前記重み  $W(I)$  を決定するステップと、
  - c) 該群の前記  $W(I)$  個の関連付けられた入力記号を決定するステップとを含む、ステップと、
- 2) 任意の入力記号を復号化するのに十分な情報が受信されたかどうかを決定するステップと、
- 3) 該受信された情報から復号化され得る入力記号を復号化するステップと、を含む、請求項 19 に記載の方法。

10

【請求項 21】

前記キー I を決定するステップが、前記パケット通信チャネルを介して受信されたパケットにおいて供給されたデータから該キー I を少なくとも部分的に決定するステップを含む、請求項 20 に記載の方法。

【請求項 22】

前記データを復号化するステップが、

- 1) 受信された出力記号の各群を処理するステップであって、
  - a) 該受信された群の前記重み  $W(I)$  を決定するステップと、
  - b) 該受信された群の前記  $W(I)$  個の関連付けられた入力記号を決定するステップと、
  - c) 出力記号の群のテーブルにおける群の出力記号の値  $B(I)$  を、該受信された群に対する該重み  $W(I)$  および前記  $W(I)$  個の関連付けの位置とともに格納するステップとを含む、ステップと、
- 2) 出力記号のさらなる群を受信し、それらをステップ 1) およびそのサブステップにしたがって処理するステップと、
- 3) 多くとも、群サイズの重みを有し、使い古された出力記号の群として表記されていない出力記号の各群、GOS1、に対して、以下のステップ、
  - a) GOS1 に対応する入力記号の位置に対して入力記号を計算するステップと、

20

- b) 該出力記号の群のテーブルにおける接続された出力記号の群を識別するステップであって、接続された出力記号の群は、ステップ 3) a) において処理された少なくとも 1 つの入力記号の関数である出力記号の群である、ステップと、
  - c) ステップ 3) a) において処理された該入力記号に独立になるように該接続された出力記号の群を再計算するステップと、
  - d) ステップ 3) a) において処理された該入力記号の独立性を反映するようにステップ 3) c) において再計算された該出力記号の群の重みを決定するステップと、
  - e) GOS1 を使い古した出力記号として表記するステップと、を実行するステップと、

30

- 4) 前記入力記号の順列セットが前記目的地で回復されるまで、該ステップ 1) からステップ 3) を繰り返すステップと、を含む、請求項 19 に記載の方法。

40

【請求項 23】

前記表記するステップが、前記使い古した出力記号にゼロの重みを割り当てるステップである、請求項 22 に記載の方法。

【請求項 24】

前記表記するステップが、前記出力記号の群のテーブルから前記使い古した出力記号の群を除去するステップを含む、請求項 22 に記載の方法。

【請求項 25】

前記パケット化するステップが、複数の出力記号の群のうちの少なくとも 1 つの群から少なくとも 1 つの出力記号の群を各パケットにパケット化するステップであり、該方法が

50

、パケット内の出力記号の 1 群の位置を該出力記号の群のキーの一部として使用するステップをさらに含む、請求項 19 に記載の方法。

【請求項 26】

前記計算するステップが、前記群における出力記号の数  $G(I)$  を計算するステップを含み、該  $G(I)$  は、正の整数であり、 $I$  のすべての値に対して同じ正の整数である、請求項 1 に記載の方法。

【請求項 27】

出力記号のグループを生成する方法であって、グループは 1 個以上の出力記号であり、出力記号の各々は出力アルファベットから選択され、該グループは、各々が入力アルファベットから選択される順列の複数の入力記号を含む入力ファイルがこのようなグループのセットから回復可能であるようなものであり、該方法は、

所与のグループに対して、該グループに関連付けられる  $W$  個の関連付け入力記号を指示するリスト  $AL$  を決定するステップであって、 $W$  は正の整数であり、少なくとも 2 つのグループは各々の  $W$  個の関連付けに対して異なる値を有し、少なくとも 1 つのグループに対して  $W$  は 2 以上であり、生成可能なグループの個数は、任意の予想される入力ファイルサイズについて、該入力ファイル内の入力記号の数に対して実質的に無限である、ステップと、

$AL$  により指示される  $W$  個の関連付け入力記号を 変数とする 所定の関数から、前記グループ内の各々の出力記号の出力記号値を生成するステップと、  
を含むことを特徴とする方法。

【請求項 28】

さらに、

所与のグループに関連付けられたキー  $I$  から  $W$  を計算するステップと、

キー  $I$  から  $AL$  を計算するステップと、

を含むことを特徴とする請求項 27 に記載の方法。

【請求項 29】

前記  $W$  を計算するステップと前記  $AL$  を計算するステップとは、確率分布を使用することを特徴とする請求項 28 に記載の方法。

【請求項 30】

入力ファイルの情報コンテンツを送信する方法であって、該入力ファイルは、各々が入力アルファベットから選択される順列の複数の入力記号であり、該入力ファイルは、各々が出力アルファベットから選択される出力記号のセットから回復可能なものであり、該方法は、

出力記号の複数のグループを生成するステップであって、グループは 1 以上の出力記号であり、前記複数のグループの各々のグループは、

a) 前記グループに関連付けられた  $W$  個の関連付け入力記号を指示するリスト  $AL$  を決定するステップであって、 $W$  は正の整数であり、少なくとも 2 つのグループは各々の  $W$  個の関連付けに対して異なる値を有し、少なくとも 1 つのグループに対して  $W$  は 2 以上であり、生成可能なグループの個数は、任意の予想される入力ファイルサイズについて、入力ファイル内の入力記号の数に対して、実質的に無限である、ステップと、

b)  $AL$  により指示される  $W$  個の関連付け入力記号を 変数とする 所定の関数から、前記グループ内の各々の出力記号の出力記号値を生成するステップと、  
により生成される、ステップと、

生成された複数のグループを送信するステップと、  
を含むことを特徴とする方法。

【請求項 31】

前記複数のグループを生成するステップは、複数のソースの各々において、独立した複数のグループを生成するステップであり、該複数のグループは、前記複数のソースの各々のソースが、該複数のソースの他のソースで生成されたりリストの参照を必要とすること無く、ある独立ソースが他のソースの情報出力と完全に重複すること無く、関連付け入

カシンボルのリストを決定するという点で独立であることを特徴とする請求項 30 に記載の方法。

【請求項 32】

さらに、

あるチャンネルにより提供される情報コンテンツが、他のチャンネルの情報コンテンツと完全に重複すること無く、複数のチャンネルにわたって複数のグループを配信するステップを含むことを特徴とする請求項 30 に記載の方法。

【請求項 33】

前記決定するステップは、

キーアルファベットから、生成されるグループに対するキー  $I$  を選択するステップであって、該キーアルファベットは入力記号の順列のセットに含まれる入力記号の個数よりも多い個数を含む、ステップと、

$I$  の関数として重み  $W(I)$  を決定するステップと、

$I$  の関数に従って入力記号の  $W(I)$  を選択ステップであって、それにより前記グループに関連付けられた  $W(I)$  個の入力記号のリスト  $AL(I)$  を形成する、ステップと、前記グループ内の出力記号の個数  $G(I)$  を決定するステップと、

関連付けられた  $W(I)$  個の入力記号を 変数とする 所定の値関数から、前記グループの  $G(I)$  個の出力記号の各々の値  $B(I)$  を計算するステップと、を含むことを特徴とする請求項 30 に記載の方法。

【請求項 34】

前記複数のグループを生成するステップは、複数のソースの各々において、独立した複数個のグループを生成するステップであり、該複数個のグループは、前記複数のソースの各々のソースが、該複数のソースの他のソースで生成されたキーの参照を必要とすること無く、ある独立ソースが他のソースの情報出力と完全に重複すること無く、キーを選択するという点で独立であることを特徴とする請求項 33 に記載の方法。

【請求項 35】

さらに、

あるチャンネルにより提供される情報コンテンツが、他のチャンネルの情報コンテンツと完全に重複すること無く、複数のチャンネルにわたって複数のグループを配信するステップを含むことを特徴とする請求項 33 に記載の方法。

【請求項 36】

複数のチャンネルを介してソースからデータを送信する方法であって、該方法は、

a) 順列の入力記号のセットとして送信されるデータを編成するステップと、

b) 複数のチャンネルの各々に対して出力記号の複数のグループを生成するステップであって、該複数のグループの各々のグループは、

1) キーアルファベットから、生成されるグループに対するキー  $I$  を選択するステップであって、該キーアルファベットは入力記号の順列のセットに含まれる入力記号の個数よりも多い個数を含む、ステップと、

2)  $I$  の関数として重み  $W(I)$  を決定するステップであって、重み  $W(I)$  は、少なくとも 2 つの値の間で変化する正の整数であり、前記キーアルファベットにおける少なくともいくつかのキーに 関して重み  $W(I)$  の値が、1 より大きい、ステップと、

3)  $I$  の関数に従って入力記号の  $W(I)$  を選択するステップであって、それにより前記グループに関連付けられた  $W(I)$  個の入力記号のリスト  $AL(I)$  を形成する、ステップと、

4) 前記グループの出力記号の個数  $G(I)$  を決定するステップと、

5) 関連付けられた  $W(I)$  個の入力記号を 変数とする 所定の値関数から、前記グループの  $G(I)$  個の出力記号の各々の値  $B(I)$  を計算するステップと、により生成される、ステップと、

c) 前記複数のグループの少なくとも 1 つを複数のパケットの各々にパケット化するステップと、

d) 前記複数のチャンネルにわたって前記複数のパケットを送信するステップであって、該複数のチャンネルの少なくとも2つは、キーIに対して互いに異なる値として生成されたグループを含むパケットを運ぶ、ステップと、  
を含むことを特徴とする方法。

【請求項37】

前記ソースは、自ソースのキーIに対する値を生成する単一のソースを含み、前記複数のチャンネルの任意の2つから受信されるグループが完全な重複とはならないように、前記複数のチャンネルにわたって結果のグループを配信する、ことを特徴とする請求項36に記載の方法。

【請求項38】

前記複数のチャンネルの任意の2つから受信されるグループが完全な重複とはならないように、キーIに対する前記値は、前記複数のチャンネルの各々に対して独立に選択されることを特徴とする請求項36に記載の方法。

【発明の詳細な説明】

(関連出願への相互参照)

本出願は、発明の名称が「INFORMATION ADDITIVE CODE GENERATOR AND DECODER FOR COMMUNICATION SYSTEMS」で1999年2月5日に出願された米国特許出願第09/246,015号(以下、「Luby I」と称す)の一部継続である米国特許出願第09/399,201号に基づく優先権を主張する。米国特許出願第09/399,201号はまた、発明の名称が「ENCODING AND DECODING DATA IN COMMUNICATION SYSTEMS」であり1998年9月23日に出願された同時係属中の米国仮特許出願第60/101,473号(以下、「仮Luby」と称す)に基づく優先権を主張する。本出願は、出力記号表現のベースとしてのリードソロモン符号などの群符号の使用をLuby Iに示し得たものよりも詳細に例示する。

【0001】

(発明の背景)

本発明は、通信システムにおけるデータの符号化および復号化に関し、より詳細には、通信データにおける誤りおよびギャップを分析し、1つより多くのソースから生じる通信データを効率的に利用するためのデータを符号化および復号化する通信システムに関する。

【0002】

送信者と受信者との間で通信チャンネルを介したファイルの送信は、多くの文献に取り上げられてきた。好ましくは、受信者は、あるレベルの確実度で送信者によってチャンネルを介して送信されたデータの正確なコピーを受信することを望む。チャンネルが完全な忠実度を有さない場合(ほとんどすべての物理的に実現可能なシステムに当てはまる)の1つの問題点は、送信中に損失したりまたは化けたデータをどのように扱うかである。損失データ(イレーズ(erasure))は、化けたデータ(誤り)よりも扱いやすいことが多い。なぜなら、受信者は、いつ化けたデータが誤ってデータ受信されたのかを必ずしも当てることができないからである。イレーズを補正するため(いわゆる「イレーズ符号」)および/または誤りを補正するため(「誤り補正符号」または「ECC」)の多くの誤り補正符号が開発されてきた。一般に、データの送信中のチャンネルおよびその送信されるデータの性質の不忠実度についての何らかの情報に基づいて、使用される特定の符号が選択される。例えば、チャンネルが長時間の不忠実度を有することが既知の場合、そのようなアプリケーションには、バースト誤り符号が最も適切であり得る。短くまれな誤りだけが予想される場合、簡単なパリティ符号が最善であり得る。

【0003】

複数の送信者および/または複数の受信器との間の通信チャンネルを介したファイル送信もまた、多くの文献に取り上げられてきた。一般に、複数の送信者からのファイル送信は、送信者が作業の重複を最小限にできるように複数の送信者の間で調整が必要である。1つのファイルを受信器に送信する一般の複数送信者システムにおいて、送信者がどのデータ

10

20

30

40

50

をいつ送信するかを調整せずにその代わりにファイルのセグメントをただ送るだけの場合は、受信器が無駄な重複セグメントを多く受信する可能性がある。同様に、一人の送信者の送信に異なる受信器が異なる時点で参加する場合の問題点は、どのようにすればその送信者から受信器が受信するすべてのデータが確実に有用となるかである。例えば、送信者が連続して同じファイルについてのデータを送信しているとする。送信者が元のファイルのセグメントを送信し、いくつかのセグメントが損失した場合、受信器は、そのファイルの各セグメントの1コピーを受信する前に多くの無駄な重複セグメントを受信し得る。

#### 【0004】

符号を選択する際に別に考慮することは、送信に使用するプロトコルである。「インターネット」(大文字の「I」を使用)として知られるネットワークのグローバルなインターネットワークの場合、パケットプロトコルがデータ伝送のために使用される。そのプロトコルは、インターネットプロトコル、または略して「IP」と呼ばれる。1ファイルまたはデータの他のブロックがIPネットワークを介して送信される場合、そのデータは、等しいサイズの入力記号に分割され、入力記号は連続パケットに配置される。パケットベースであるので、パケット指向符号化スキームが適切であり得る。入力記号の「サイズ」は、入力記号が実際にビットストリームに分割されるどうかに関わらず、ビット単位で測られ得る。この場合、入力記号が $2^M$ 記号のアルファベットから選択されるとすると、入力記号の大きさは、Mビットのサイズを有する。

#### 【0005】

伝送制御プロトコル(「TCP」)は、確認応答機構を有する常用のポイントツーポイントパケット制御スキームである。TCPは、一対一通信に適切である。ここで、送信者および受信者の双方は、いつ送信および受信が生じるかについての合意があり、およびどの送信器および受信器が使用されるかの合意がある。しかし、TCPは、一対多または多対多通信または送信者および受信者が独立してデータの送信または受信の時間と場所を決定する場合に対しては、適切でないことがよくある。

#### 【0006】

TCPを使用する場合、送信者は整列したパケットを送信し、受信者は各パケットの受信の確認応答を行う。パケットが損失すると、送信者に確認応答が送られないので、送信者はパケットを再送信する。パケット損失は、多くの原因による。インターネット上では、パケット損失が頻繁に起こる。なぜなら、散発的な混雑状態によってルータにおけるバッファリング機構がその容量の最大に達し、入力パケットの損失を起こすからである。TCP/IPなどのプロトコル使用すると、確認応答パラダイムにより、確認応答の欠如に回答して、または受信者からの明示的要求に回答して損失パケットが単に再送信され得るので、パケット送信は完全には失敗しない。いずれの場合も、確認応答プロトコルは、受信者から送信者の逆向きチャネルを必要とする。

#### 【0007】

確認応答ベースプロトコルは、一般に多くのアプリケーションに適切であり、実際に現在のインターネットにおいて広く使用されているが、あるアプリケーションに対しては、それらは非効率的であり、完全に実行不可能であることもある。特に、確認応答ベースプロトコルは、待ち時間が長く、パケット損失率が高く、調整されていない受信者の接続および切断があり、および/または帯域の非対称性が高いようなネットワーク中では、十分に機能しない。長い待ち時間とは、確認応答が受信者から伝送して送信者に戻るまでに要する時間が長いことである。待ち時間が長いと、再送信前の総時間が過度に長くなり得る。パケット損失率が高くなるとまた、同じパケットのうちのそれぞれの再送信が到達せずに最後の1つまたは最後の数個の不運なパケットを得ることが大きく遅れるという問題を起こす。

#### 【0008】

「調整されていない受信者の接続および切断」は、各受信者が進行中の送信セッションに接続および切断を自由裁量でできる状況のことである。このような状況は、インターネット、「ビデオ・オン・デマンド」などの次世代サービスおよび将来のネットワークプロバ

10

20

30

40

50

イダによって提供される他のサービス上では一般的なことである。一般のシステムにおいて、受信者は、送信者の調整なしに進行中の送信に接続および切断する場合、受信者が多数のパケットの損失を認識する可能性がある。この場合、異なる受信者によって広く異なる損失パターンが認識される。

#### 【 0 0 0 9 】

非対称性帯域は、受信者から送信者への逆向きデータパス（逆向きチャネル）が利用できないか、または前向きパスよりもさらにコストがかかる状況のことである。非対称性帯域の場合、受信者がパケットに頻繁に確認応答するのを過度に遅くさせ、および／または過度にコストがかかる。また、確認応答を頻繁にしなければやはり遅延が生じる。

#### 【 0 0 1 0 】

さらに、確認応答ベースプロトコルは、放送にまで規模を十分に拡大することができない。放送では、1送信者が複数のユーザに同時にファイルを送信する。例えば、1送信者が複数の受信者に衛星チャネルを介してファイルを放送すると想定する。各受信者は、異なるパターンでパケットが損失し得る。ファイルの確実な送達のために確認応答データ（肯定または否定のいずれか）に基づくプロトコルは、各受信者から送信者への逆向きチャネルを必要とし、それを提供するのは過度にコストがかかり得る。さらに、このような場合には、複雑で強力な送信者が受信者から送信された確認応答データのすべてを適切に処理することができる必要がある。別の欠点は、異なる受信者が異なるパケットのセットを損失する場合に、ほんの少数の受信者によって受信し損ねたパケットのみを再放送することにより、他の受信者が無駄な重複パケットを受信してしまうことである。確認応答ベース通信システムにおいて十分に扱えない別の状況は、受信者が非同期的に受信セッションを開始し得る場合であり、つまり、受信者が送信セッションの途中でデータ受信を開始する可能性のある場合である。

#### 【 0 0 1 1 】

マルチキャストおよび放送のためのTCP/IPなどの確認応答ベーススキームの性能を向上させるための複雑なスキームがいくつか提案されてきた。しかし、現在のところ種々の理由によりいずれも完全には採用されていない。ひとつには、低地球軌道（LEO）衛星放送ネットワークなどのような1受信者が複数の送信者から情報を得るような場合にまで確認応答ベースプロトコルの規模を十分に拡大することができないことである。LEOネットワークにおいては、LEO衛星はその軌道のために上空を速い速度で通過するので、受信者は、いずれの特定の衛星の視野にはほんの短時間しか存在しない。これを補うために、LEOネットワークは多くの衛星を含み、1つの人工衛星が地平線に沈み別の衛星が地平線を昇る際に受信者が衛星間で引き渡される。確認応答ベースプロトコルが確実性を確保するために使用されとしても、確認応答が適切な衛星から戻るように調整するためには、おそらく複雑な引き渡しプロトコルが必要であろう。なぜなら、受信者が1つの衛星からパケットを受信しつつなお別の衛星へそのパケットの確認応答を送ることがよく起こり得るからである。

#### 【 0 0 1 2 】

確認応答ベースプロトコルを代替する、実用されることのあるプロトコルは、カルーセル（carousel）ベースプロトコルである。カルーセルベースプロトコルは、入力ファイルを等しい長さの入力記号に分割し、各入力記号をパケットに配置し、そして連続的にサイクルを繰り返す、すべてのパケットを送信する。カルーセルベースプロトコルの主な欠点は、受信者がたった1つのパケットを受信し損ねても、受信者がさらに完全な1回のサイクルを待たなければその受信し損ねたパケットを受信する機会を得られないことである。別の観点では、カルーセルベースプロトコルのために大量の無駄な重複データの受信を生じ得ることである。例えば、受信者がカルーセルの開始からパケットを受信し、しばらく受信を停止し、カルーセルの開始時に受信を再開する場合、多くの無駄な重複パケットが受信される。

#### 【 0 0 1 3 】

上記問題を解決するために提案されてきた1つの解決法は、確認応答ベースプロトコルの

10

20

30

40

50

使用を避け、その代わりに確実性を上げるためにリードソロモン符号などのイレース符号を使用することである。いくつかのイレース符号の1つの特徴は、ファイルが入力記号にセグメント化され、その入力記号がパケットで受信者に送信される場合、受信者は、十分に多くのパケットを受信したら、一般にどのパケットが到着したかにかかわらず、そのパケットを復号して完全なファイルを再構築する。このような性質は、たとえパケットが損失してもファイルが回復され得るので、パケットレベルでの確認応答の必要性を排除する。しかし、イレース符号による多くの解決方法は、確認応答ベースプロトコルの問題を解決しないか、さもなければ新規の問題を生じる。

#### 【0014】

多数のイレース符号に伴う1つの問題は、それら进行处理するために過剰な計算能力またはメモリを必要とすることである。計算能力およびメモリを多少効率的に使用する通信アプリケーション用に最近開発された1つの符号化スキームは、トルネード符号化スキームである。トルネード符号は、入力ファイルがK個の入力記号によって表されN個の出力記号を決定するために使用される(ここでNは、符号化処理を開始する前に固定される)という点で、リードソロモン符号と同様である。トルネード符号を使用する符号化は、一般にリードソロモン符号を使用する符号化よりもずっと速い。なぜなら、N個のトルネード出力記号を生成するのに必要な平均算術演算数がNに比例し(数十アセンブリ符号演算かけるNのオーダーである)、ファイル全体を復号化するのに必要な算術演算総数がまたNに比例するからである。

#### 【0015】

トルネード符号は、リードソロモン符号よりも速度の点で優れているが、いくつかの欠点がある。まず、出力記号数Nが符号化処理の前に決定されなければならない。このため、パケットの損失率が過大に見積もられると効率が低下し、パケットの損失率が過小に見積もられると失敗する。これは、トルネードデコーダが、元のファイルを復号して復元するために所定数の出力記号(具体的には、 $K + A$ 個の出力記号、ここでAはKに比較して小さい)を必要とし、損失出力記号数が $N - (K + A)$ よりも大きい場合、元のファイルが復元されないからである。この制限は、一般にNが $K + A$ よりも少なくとも実際のパケット損失分は大きくなるように選択される限り、多くの通信の問題に対して許容されるが、パケット損失をあらかじめ予測する必要がある。

#### 【0016】

トルネード符号の別の欠点は、トルネード符号がグラフ構造上で何らかの方法で整合するためにエンコーダおよびデコーダを必要とすることである。トルネード符号は、このグラフが構築されるデコーダで前処理段階を必要とする。その処理は、実質的に復号化を遅らす。さらに、グラフはファイルサイズに固有であるので、使用される各ファイルサイズに対する新しいグラフを生成する必要がある。さらに、トルネード符号によって必要なグラフの構築は複雑であり、最高の性能を得るためには、異なるサイズのファイルに対するパラメータの異なるカスタム設定をする必要がある。これらのグラフは、大きなサイズを有し、送信者および受信者の双方の記憶装置に大容量のメモリを要求する。

#### 【0017】

加えて、トルネード符号は、固定されたグラフおよび入力ファイルに対してちょうど同じ出力記号値を生成する。これらの出力記号は、K個の元入力記号および $N - K$ 個の冗長記号を含む。さらに実用上、Nは単にKの小さな倍数である(Kの1.5または2倍など)。このように、1より多い送信者から、同じグラフを使用して同じ入力ファイルから生成された出力記号を得る受信者が、多数の無駄な重複出力記号を受信する可能性は非常に高い。なぜなら、N個の出力記号があらかじめ固定されており、そのN個の出力記号は、出力記号が送信されるごとの各送信器から送信されるN個の出力記号と同じであり、かつ受信者によって受信されるN個の記号と同じである。例えば、 $N = 1500$ 、 $K = 1000$ 、および受信者が、1つの衛星が地平線へ沈む前にその衛星から900個の記号を受信すると想定する。衛星が同期するように調整されなければ、次の衛星から受信器によって受信されたトルネード符号記号は、追加のものではない。なぜなら、その次の衛星は、同じ

10

20

30

40

50

N個の記号を送信しているからである。この場合、受信器は、入力ファイルの復元に必要な100個の新しい記号を受信する前にすでに受信した900個の記号のうちの多く記号のコピーを受信してしまう可能性がある。

【0018】

したがって、必要なことは、送信者または受信者に過剰な計算能力またはメモリを設置する必要のない、かつ必ずしも1以上の送信者と1以上の受信者との間の調整を必要とせずにその送信者および/またはその受信者によってシステム内にファイルを効率的に配布するために使用され得る簡単なイレース符号である。

【0019】

(発明の要旨)

本発明は以下を提供する。

【0020】

1. 出力記号の1群を生成する方法であって、1群は1つ以上の出力記号であり、各出力記号は出力アルファベットから選択され、入力アルファベットからそれぞれ選択される順列の複数の入力記号を含む入力ファイルが、このような群の1セットから回復可能であるような該群であり、該方法は、該群のキーIを得るステップであって、該キーがキーアルファベットから選択され、該キーアルファベットにおける可能なキーの数が該入力ファイルにおける入力記号の数よりもずっと大きい、ステップと、Iの所定の関数にしたがって、該群のリストAL(I)を計算するステップであって、AL(I)は、該群に関連付けられたW(I)個の関連付けられた入力記号を示し、重みWは、少なくとも2つの値の間で変化する正の整数であり、Iの少なくとも1つの値に対して1より大きい、ステップと、AL(I)によって示される該関連付けられた入力記号の所定の関数から、該群における各出力記号の出力記号値を生成するステップと、を含む方法。

【0021】

2. 前記キーIを得るステップが、ランダム関数または擬似ランダム関数にしたがってキーIを計算するステップを含む、項目1に記載の方法。

【0022】

3. 前記計算するステップが、Iのランダム関数または擬似ランダム関数にしたがってW(I)を計算するステップを含む、項目1に記載の方法。

【0023】

4. 前記計算するステップが、Iのランダム関数または擬似ランダム関数にしたがってAL(I)を計算するステップを含む、項目1に記載の方法。

【0024】

5. 前記計算するステップが、前記群における出力記号の数G(I)を計算するステップを含み、該数G(I)が正の整数であり、該数G(I)がIのランダム関数または擬似ランダム関数にしたがって計算される、項目1に記載の方法。

【0025】

6. 前記計算するステップが、Iの所定の関数および確率分布にしたがって、重みW(I)を計算するステップであって、該確率分布が少なくとも2つの正の整数上にあり、そのうちの少なくとも1つが1より大きい、ステップと、リストAL(I)のリストエントリを計算するステップと、W(I)個のリストエントリが計算されるまで、リストAL(I)のリストエントリを計算するステップを繰り返すステップと、を含む、項目1に記載の方法。

【0026】

7. 前記W(I)を計算するステップが、Wが前記キーアルファベット上の所定の分布を近似するようなW(I)を決定するステップを含む、項目6に記載の方法。

【0027】

8. 前記所定の分布が一様分布である、項目7に記載の方法。

【0028】

9. 前記所定の分布が釣り鐘曲線分布である、項目7に記載の方法。

10

20

30

40

50

## 【 0 0 2 9 】

1 0 . 前記所定の分布は、 $W = 1$  が  $1 / K$  の確率を有し、ここで、 $K$  が前記入力ファイルにおける入力記号の数であり、そして  $W = i$  が  $i = 2, \dots, K$  に対して  $1 / i (i - 1)$  の確率を有する、項目 7 に記載の方法。

## 【 0 0 3 0 】

1 1 . 前記  $AL(I)$  によって示される関連付けられた入力記号の前記所定の関数が、リード - ソロモン符号によって決定される、項目 1 に記載の方法。

## 【 0 0 3 1 】

1 2 . 前記入力アルファベットおよび前記出力アルファベットが同じアルファベットである、項目 1 に記載の方法。

## 【 0 0 3 2 】

1 3 . 前記入力アルファベットが  $2^{M_i}$  個の記号を含み、各入力記号が  $M_i$  ビットを符号化し、前記出力アルファベットが  $2^{M_o}$  個の記号を含み、各出力記号の群が  $M$  ビットを符号化する、項目 1 に記載の方法。

## 【 0 0 3 3 】

1 4 . 各後続キー  $I$  がその先行キーよりも 1 大きい、項目 1 に記載の方法。

## 【 0 0 3 4 】

1 5 . 出力記号の複数の群を符号化する方法であって、各群は項目 1 に記載され、該方法は、生成されるべき該出力記号の群の各々に対するキー  $I$  を生成するステップと、および該生成された出力記号の群の各々をデータレイズチャネルを介して送信されるべき出力シーケンスとして出力するステップと、をさらに含む方法。

## 【 0 0 3 5 】

1 6 . 各キー  $I$  が他の選択されるキーとは独立に選択される、項目 1 5 に記載の方法。

## 【 0 0 3 6 】

1 7 . 前記  $AL(I)$  を計算するステップが、前記入力ファイルにおける入力記号の数  $K$  を少なくとも近似的に識別し、かつ重み  $W(I)$  を識別するステップと、 $K$  より大きいまたは  $K$  に等しい最小素数  $P$  を決定するステップと、 $P$  が  $K$  より大きい場合、 $P - K$  個のパディング入力記号で該入力ファイルを少なくとも論理的にパディングするステップと、 $1 \leq X < P$  である第 1 の整数  $X$ 、および  $0 \leq Y < P$  である第 2 の整数  $Y$  を生成するステップと、 $1 \sim W(I)$  の各  $J$  に対して、 $AL(I)$  における第  $J$  番目のエントリを  $((Y + (J - 1) \cdot X) \bmod P)$  に設定するステップと、を含む、項目 1 に記載の方法。

## 【 0 0 3 7 】

1 8 . 前記各  $J$  に対して  $AL(I)$  における第  $J$  番目のエントリを設定するステップが、配列  $V$  における第 1 のエントリ  $V[J = 0]$  を  $Y$  に設定するステップと、 $1 \sim W(J) - 1$  の各  $J$  に対して、該配列  $V$  における第  $J$  番目のエントリ  $V[J]$  を  $((V[J - 1] + X) \bmod P)$  に設定するステップと、該配列  $V$  を該リスト  $AL(I)$  として使用するステップと、を含む、項目 1 7 に記載の方法。

## 【 0 0 3 8 】

1 9 . ソースから目的地にパケット通信チャネルを介してデータを送信する方法であって、該方法が、a) 送信されるべき該データを入力記号の順列セットとして構成するステップであって、各入力記号は、入力アルファベットから選択され、該データにおける位置を有する、ステップと、b) 出力記号の複数の群を生成するステップであって、各出力記号が出力アルファベットから選択され、該複数の群の各群が、以下のステップによって生成され、該ステップは、キーアルファベットから、生成される該群のキー  $I$  を選択するステップと、 $I$  の関数として重み  $W(I)$  を決定するステップであって、ここで、重み  $W$  は、少なくとも 2 つの値の間および該キーアルファベット上で変化する正の整数であり、該キーアルファベットにおける少なくともいくつかのキーに対しては 0 より大きい、ステップと、 $I$  の関数にしたがって  $W(I)$  個の該入力記号を選択し、したがって、該群に関連付けられる  $W(I)$  個の入力記号のリスト  $AL(I)$  を形成するステップと、該群における

10

20

30

40

50

出力記号の数  $G(I)$  を決定するステップと、該関連付けられる  $W(I)$  個の入力記号の所定の値関数から該群における各出力記号の値  $B(I)$  を計算するステップとを含む、ステップと、c) 該複数の群のうちの少なくとも1つの出力記号の群のうちの少なくとも1つを、複数のパケットの各々にパケット化するステップと、d) 該複数のパケットを該パケット通信チャネルを介して送信するステップと、e) 該複数のパケットのうちの少なくともいくつかを該目的地で受信するステップと、f) 該複数の受信されたパケットから該データを復号化するステップと、を含む方法。

#### 【0039】

20. 前記データを復号化するステップが、1) 受信された出力記号の各群を処理するステップであって、a) 該群の前記キー  $I$  を決定するステップと、b) 該群の前記重み  $W(I)$  を決定するステップと、c) 該群の前記  $W(I)$  個の関連付けられた入力記号を決定するステップとを含む、ステップと、2) 任意の入力記号を復号化するのに十分な情報が受信されたかどうかを決定するステップと、3) 該受信された情報から復号化され得る入力記号を復号化するステップと、を含む、項目19に記載の方法。

10

#### 【0040】

21. 前記キー  $I$  を決定するステップが、前記パケット通信チャネルを介して受信されたパケットにおいて供給されたデータから該キー  $I$  を少なくとも部分的に決定するステップを含む、項目20に記載の方法。

#### 【0041】

22. 前記データを復号化するステップが、1) 受信された出力記号の各群を処理するステップであって、a) 該受信された群の前記重み  $W(I)$  を決定するステップと、b) 該受信された群の前記  $W(I)$  個の関連付けられた入力記号を決定するステップと、c) 出力記号の群のテーブルにおける群の出力記号の値  $B(I)$  を、該受信された群に対する該重み  $W(I)$  および前記  $W(I)$  個の関連付けの位置とともに格納するステップとを含む、ステップと、2) 出力記号のさらなる群を受信し、それらをステップ1) およびそのサブステップにしたがって処理するステップと、3) たかだか群サイズの重みを有し、使い古された出力記号の群として表記されていない出力記号の各群、 $GOS1$ 、に対して、以下のステップ、a)  $GOS1$  に対応する入力記号の位置に対して入力記号を計算するステップと、b) 該出力記号の群のテーブルにおける接続された出力記号の群を識別するステップであって、接続された出力記号の群は、ステップ3) a) において処理された少なくとも1つの入力記号の関数である出力記号の群である、ステップと、c) ステップ3) a) において処理された該入力記号に独立になるように該接続された出力記号の群を再計算するステップと、d) ステップ3) a) において処理された該入力記号の独立性を反映するようにステップ3) c) において再計算された該出力記号の群の重みを決定するステップと、e)  $GOS1$  を使い古した出力記号として表記するステップと、を実行するステップと、4) 前記入力記号の順列セットが前記目的地で回復されるまで、該ステップ1) からステップ3) を繰り返すステップと、を含む、項目19に記載の方法。

20

30

#### 【0042】

23. 前記表記するステップが、前記使い古した出力記号にゼロの重みを割り当てるステップである、項目22に記載の方法。

40

#### 【0043】

24. 前記表記するステップが、前記出力記号の群のテーブルから前記使い古した出力記号の群を除去するステップを含む、項目22に記載の方法。

#### 【0044】

25. 前記パケット化するステップが、複数の出力記号の群のうちの少なくとも1つの群から少なくとも1つの出力記号の群を各パケットにパケット化するステップであり、該方法が、パケット内の出力記号の1群の位置を該出力記号の群のキーの一部として使用するステップをさらに含む、項目19に記載の方法。

#### 【0045】

26. 前記計算するステップが、前記群における出力記号の数  $G(I)$  を計算するステッ

50

プを含み、該  $G(I)$  は、正の整数であり、 $I$  のすべての値に対して同じ正の整数である、項目 1 に記載の方法。

#### 【0046】

本発明による通信システムの 1 つの実施形態において、エンコーダは、入力ファイルのデータおよびキーを使用して、出力記号の 1 群を生成する。ここで、入力ファイルは、入力アルファベットからそれぞれ選択される順列の複数の入力記号であり、キーは、キーアルファベットから選択され、群サイズは、正の整数として選択され、各出力記号は、出力アルファベットから選択される。キー  $I$  を有する出力記号の 1 群は、生成されるべき出力記号の群の群サイズ  $G(I)$  を決定するステップ（ここで、群サイズ  $G$  は、複数のキー上の少なくとも 2 つの値の間で変化し得る正の整数である）、生成されるべき出力記号の群の重み  $W(I)$  を決定するステップ（ここで、重み  $W$  は、複数のキー上の少なくとも 2 つの値の間で変化する正の整数である）、 $I$  の関数にしたがって、出力記号の群に関連付けられる  $W(I)$  個の入力記号を選択するステップ、および選択された  $W(I)$  個の入力記号の所定の値関数  $F(I)$  から  $G(I)$  個の出力記号値の群  $B(I) = B\_1(I), \dots, B\_G(I)(I)$  を生成するステップによって生成される。エンコーダは、1 回以上呼び出され得、各回別のキーを用い、各回出力記号の 1 群を生成する。出力記号の群は、一般に互いに独立であり、無限の数（ $I$  の解像度に従う）が、必要ならば生成され得る。

10

#### 【0047】

いくつかの実施形態において、1 群における出力記号数  $G(I)$  は、キー  $I$  に依存して群ごとに異なる。他の実施形態においては、 $G(I)$  は、すべてのキー  $I$  の出力記号のすべての群に対して、何らかの正の整数  $b$  に固定される。

20

#### 【0048】

本発明によるデコーダにおいて、受信者によって受信された出力記号は、入力ファイルの符号化に基づいてこれらの出力記号を群単位で生成した送信者から送信された出力記号である。出力記号は送信中に損失され得るが、デコーダは、送信された出力記号の任意の部分を受信するだけの場合でさえ適切に動作する。入力ファイルを復号化するために必要な出力記号数は、そのファイルを構成する入力記号の数に等しいか、またはわずかに大きい。ここで、入力記号および出力記号は、同じビット数のデータと仮定する。

#### 【0049】

本発明による 1 つの復号化プロセスにおいて、出力記号の各受信された群に対して以下のステップが実行される：1) 受信された  $b$  個の出力記号からなる群のキー  $I$  を識別するステップ；2) 群サイズ  $G(I)$  を識別するステップ；3) 出力記号の群に対して、受信された出力記号値の群  $B(I) = B\_1(I), \dots, B\_G(I)(I)$  を識別するステップ；3) 出力記号の群の重み  $W(I)$  を決定するステップ；4) 出力記号の群に関連付けられた  $W(I)$  個の関連付けられた入力記号の位置を決定するステップ；および 5) 出力記号の群を保持するテーブルに  $B(I) = B\_1(I), \dots, B\_G(I)(I)$  を格納し、重み  $W(I)$  および関連付けられた入力記号の位置を格納するステップ。次に、以下のプロセスが、群の重みがたかだか  $G(I)$  であるように、キー  $I$  を有する出力記号の未使用群がもはや存在しなくなるまで繰り返し適用される：1)  $b' - G(I)$  の重みを有し、出力記号の「使い古した」群として表記されない、キー  $I$  を有する出力記号の各格納された群に対して、キー  $I$  に基づいて、出力記号の群に関連付けられる固有の残り  $b' - G(I)$  個の未回復入力記号の位置  $J\_1, \dots, J_{b' - G(I)}$  を計算するステップ；2) 出力記号の群から入力記号  $J\_1, \dots, J_{b' - G(I)}$  に対する未知の値、およびその群に関連付けられた他の入力記号の既知の値を計算するステップ；3) 入力記号  $J\_1, \dots, J_{b' - G(I)}$  のうちの少なくとも 1 つを関連付けとして有する出力記号の群のテーブルにおいて出力記号の群を識別するステップ；4)  $J\_1, \dots, J_{b' - G(I)}$  の中にある各関連付けに対して、これら識別された出力記号の群の各々の重みを 1 だけデクリメントするステップ；および 5) 入力記号  $J\_1, \dots, J_{b' - G(I)}$  を回復されたものとして表記し、キー  $I$  を有する出力記号の群を使い古したものとして表記するステップ。このプロセスは

30

40

50

、入力記号の順列セットが回復されるまで、すなわち、入力ファイル全体が完全に回復されるまで繰り返される。

【 0 0 5 0 】

本発明の1つの利点は、送信中に受信者が任意の所定の時点で受信を開始する必要がないこと、および出力記号の群のセット数が生成された後、送信者が停止する必要がないことである。なぜなら、送信者が任意の所定の入力ファイルに対して出力記号の群の有効な無限のセットを送信し得るためである。その代りに、受信者は、用意ができた場合に受信できるところから受信を開始し得、出力記号の群をランダムパターンでまたは任意のパターンでさえ含むパケットを損失し得、そしてなお、受信されるデータの非常に多くが「情報付加的」データ（すなわち、すでに利用可能な情報の重複でなく、回復プロセスを補助するデータ）である良好な確率を有する。独立に生成された（ランダムに無関係であることが多い）データストリームが、情報付加的な方法で符号化されることによって、複数ソース生成および受信、イレースロバスタ性、ならびに未調整の接続および切断を可能にするという多くの利点が生じる。

10

【 0 0 5 1 】

本明細書中で開示される本発明の性質および利点のさらなる理解は、本明細書の残りの部分および添付の図面を参照して理解され得る。

【 0 0 5 2 】

（好適な実施形態の説明）

本明細書中において記載される実施例において、「群連鎖反応符号化」と呼ばれる符号化スキームが説明される。そのまえに、本記載中に使用される種々の用語の意味および範囲を説明する。

20

【 0 0 5 3 】

群連鎖反応符号化を用いて、出力記号の群が必要に応じて入力ファイルから送信者によって生成される。出力記号の各群は、出力記号の他の群がどのように生成されるかにかかわらず生成され得る。どの時点においても、送信者は、出力記号の群生成工程を停止し得る。送信者が出力記号の群生成工程をいつ停止または再開するかについては、制約を必要としない。一旦生成された後、出力記号のこれらの群は、パケット内に配置され、送信先へ送信される。この場合、各パケットは、出力記号の群を1つ以上含む。群連鎖反応符号化の具体的な例では、各群は、1つの出力記号を含み、その結果、L u b y I に記載の連鎖反応符号化システムと同様の性能および動作を有するシステムを生じる。

30

【 0 0 5 4 】

本明細書中で使用されるように、用語「ファイル」は、1つ以上のソースに格納され、1つ以上の送信先へ1単位として送達される任意のデータのことである。したがって、ファイルサーバまたはコンピュータ記憶装置からのドキュメント、イメージ、およびファイルはすべて、送達され得る「ファイル」の例である。ファイルは、既知のサイズを有するか（ハードディスク上に格納された1メガバイトのイメージなど）、または未知のサイズを有する（ストリーミングソースの出力から取り出されるファイルなど）。いずれの場合も、ファイルは入力記号のシーケンスであり、各入力記号はファイル内の位置および値を有する。

40

【 0 0 5 5 】

送信は、ファイルを送達するためにデータを1以上の送信者から1以上の受信者にチャンネルを介して送信する処理のことである。1送信者が完全なチャンネルによって任意数の受信者に接続される場合、すべてのデータが正しく受信されるので受信されたデータが入力ファイルの正確なコピーであり得る。ここで、大半の現実のチャンネルの場合のようにチャンネルが完全でないと仮定するか、またはいくつかのシステムの場合のように1より多い送信者からデータが送信されると仮定するか、またはいくつかのシステムの場合のように送信されるデータの一部が途中で故意に脱落されると仮定する。多くのチャンネルの欠点のうち、ここで関心のある2つの欠点は、データイレースおよびデータ不完全性（データイレースの特殊な場合として処理され得る）である。データイレースが生じるのは、チャンネルがデー

50

データを損失または脱落させる場合である。データ不完全性が生じるのは、データをうちのいくつかを受信者を通り過ぎてしまうまで受信者がデータの受信を開始しないか、受信者が送信終了前にデータの受信を停止するか、または受信者がデータ受信の停止および再開を断続的に行う場合である。

【 0 0 5 6 】

データ不完全性の例として、移動衛星送信者が、入力ファイルを表すデータを送信し、受信者が範囲に入る前に送信を開始する。受信者が一旦範囲に入ると、衛星が範囲を出るまでデータが受信され得る。この場合、受信者は、衛星放送受信アンテナの向きを変更して（この間はデータを受信しない）、範囲内に入った別の衛星によって送信されている同じ入力ファイルについてのデータの受信を開始し得る。この記載を読むことにより明らかなように、データ不完全性は、データイレースの特殊な場合である。なぜなら、受信者は、あたかも受信者が常に範囲内に存在しているが、チャンネルは受信者がデータ受信を開始する時点までのすべてのデータを損失するかのよう、データ不完全性を処理し得る（および受信者は同じ問題を有する）。また、通信システム設計において周知のように、検出可能な誤りは、単に検出可能な誤りを有するデータブロックまたは記号のすべてを脱落させることによってイレースと等価になり得る。

10

【 0 0 5 7 】

別の例としては、ルータは、そのバッファがフルまたはフルに近い（混雑状態）場合に故意にパケットを脱落させ、かつ競合するパケットに対して公平であるようにおよび/または強制的に速度制限をするように故意にパケットを脱落させる。

20

【 0 0 5 8 】

いくつかの通信システムにおいて、受信者は、複数の送信者または複数の接続を有する 1 送信者によって生成されるデータを受信する。例えば、ダウンロードの速度を上げるために、受信者は、同時に 1 より多い送信者に接続して同じファイルに関するデータを送信する。別の例として、マルチキャスト送信において、複数のマルチキャストデータストリームは、受信者がこれらデータストリームのその 1 つ以上に接続できるよう送信され、そのストリームが送信者に接続されるチャンネルの帯域に総送信速度を整合させる。すべてのそのような場合において重要なのは、たとえ送信速度が異なるストリーム間で大きく異なる場合であっても、かつ任意の損失パターンがある場合でも、すべての送信データが受信者に対して独立に使用されること、すなわち、複数のソースデータがストリーム間で冗長でないことである。

30

【 0 0 5 9 】

一般に、送信とは、送信者から受信者へ送信者と受信者とを接続するチャンネルを介してデータを移動する行為である。そのチャンネルは、チャンネルがデータを取得した際にチャンネルがそのデータを送信者から受信者へ移動させる実時間チャンネルであり得る。あるいは、チャンネルは、送信者から受信者への送信中にデータの一部またはすべてを格納する記憶チャンネルでもよい。後者の例は、ディスク記憶装置または他の記憶装置である。この例では、データを生成するプログラムまたはデバイスは、データを記憶装置に送信する送信者として考えられ得る。受信者は、記憶装置からデータを読み出すプログラムまたはデバイスである。送信者がデータを記憶装置へ書き込むために使用する機構、記憶装置自体、および受信者が記憶装置からデータを読み出すための機構が、集合的にチャンネルを形成する。これらの機構または記憶装置がデータを損失する可能性がある場合、そのような損失をチャンネル中のデータイレースとして処理し得る。

40

【 0 0 6 0 】

送信者および受信者がデータイレースチャンネルによって分離される場合、入力ファイルの正確なコピーを送信するのではなくて、その代わりにイレースの回復を補助する入力ファイルから生成されたデータを送信するのが好ましい。エンコードとは、そのようなタスクを処理する回路、デバイス、モジュール、または符号セグメントである。エンコードの動作の見る 1 つの方法は、エンコードが入力記号から出力記号の群を生成することであり、ここで入力記号値のシーケンスが入力ファイルを表す。したがって、各入力記号は、入力

50

ファイル内の位置、および値を有する。デコーダとは、受信者によって受信された出力記号の群から入力記号を再構築する回路、デバイス、モジュール、または符号セグメントである。

#### 【 0 0 6 1 】

群連鎖反応符号化は、任意の特定のタイプの入力記号に限定されないが、入力記号のタイプはアプリケーションによって決定されることが多い。一般に、入力記号に対する値は、 $2^M$ 個の記号（ここでMは正の整数）のアルファベットから選択される。そのような場合、入力記号は、入力ファイルからのMビットのデータシーケンスによって表現され得る。Mの値は、アプリケーションの使用、チャンネル、および群の最大サイズに基づいて決定されることが多い。例えば、パケットベースインターネットチャンネルに対しては、1024 10  
バイトのサイズのペイロードを有するパケットが適切であり得る（1バイトは8ビット）。この例において、各パケットが出力記号の1群および8バイトの補助情報を含むと仮定し、かつすべての群が4個の出力記号を含むと仮定すると、入力記号サイズのMは、 $(1024 - 8) / 4$ 、すなわち254バイトが適切であり得る。別の例として、MP EGパケット規格を使用するいくつかの衛星システムがある。この場合、各パケットのペイロードは188バイトを含む。この例において、各パケットは出力記号の1群および4バイトの補助情報を含むと仮定し、かつすべての群が2個の出力記号を有すると仮定すると、記号サイズのMは、 $(188 - 4) / 2$ 、すなわち97バイトが適切であり得る。群連鎖反応符号化を使用する汎用通信システムにおいて、入力記号サイズ（すなわち、M、入力記号によって符号化されたビット数）などのアプリケーション特定パラメータは、アプリケーションによって設定された変数である。 20

#### 【 0 0 6 2 】

出力記号の各群は、その群における出力記号の各々に対する値を有する。以下を考慮する1つの好ましい実施形態において、出力記号の各群は、「キー」と呼ばれる識別子を有する。好ましくは、出力記号の各群のキーは、受信者によって容易に決定され、受信者が出力記号のある群を出力記号の別の群と区別できるようにする。好ましくは、出力記号の群のキーは、出力記号の他のすべての群のキーと異なる。また好ましくは、受信者が受信された出力記号の群のキーを決定するためには、送信中に含まれるデータは可能な限り少ない方がよい。

#### 【 0 0 6 3 】

キーイングの簡単な形態では、エンコーダによって生成された出力記号の連続した群に対するキーのシーケンスは、連続した整数のシーケンスである。この場合、各キーは、「シーケンス番号」と呼ばれる。各送信パケット中の出力記号値の1群がある場合、シーケンス番号がパケット中に含まれ得る。一般にシーケンス番号が少数のバイト（例えば、4バイト）に収まるので、いくつかのシステムにおいては、出力記号値の群とともにシーケンス番号を含むことは経済的である。例えば、各1024バイトのUDPインターネットパケットを使用して、シーケンス番号用に各パケット内に4バイトを割当ててもわずかに0.4%の負担がかかるだけである。

#### 【 0 0 6 4 】

他のシステムでは、1より多くのデータからキーを生成するのが好ましい。例えば、1以上の送信者からの同じ入力ファイルに基づいて生成された1より多くのデータストリームを受信する受信者を含むシステムを考える。ここで、送信データはパケットのストリームであり、各パケットは、出力記号の1群を含む。すべてのそのようなストリームが、キーと同じシーケンス番号のセットを使用する場合、受信者は、同じシーケンス番号を有する出力記号の群を受信する可能性がある。同じキーを有するか、または同じシーケンス番号を有する場合の出力記号の群は、入力ファイルについて同一の情報を含むので、受信者は、無駄な重複データを受信する。したがって、そのような場合、キーがシーケンス番号と対になった固有のストリーム識別子を含むことが好ましい。

#### 【 0 0 6 5 】

例えば、UDPインターネットパケットの1ストリームに対して、データストリームの固 50

有の識別子は、送信者のIPアドレス、および送信者がパケットを送信するために使用しているポート数を含み得る。送信者のIPアドレスおよびストリームのポート数は、各UDPパケットのヘッダの一部であるので、キーのこれらの部分が受信者に確実に利用され得るように各パケット中に必要なさらなるスペースはない。送信者は、単にシーケンス番号を各パケットに対応の出力記号の群とともに挿入することのみを必要とし、受信者は、シーケンス番号およびパケットヘッダから受信された出力記号の1群のキー全体を再生成し得る。別の例として、IPマルチキャストパケットの1ストリームに対して、データストリームの固有の識別子は、IPマルチキャストアドレスを含み得る。IPマルチキャストアドレスは各IPマルチキャストパケットのヘッダの一部であるので、UDPパケットについての上記説明は、この場合についても同様に適用される。

10

**【0066】**

出力記号の群の位置によるキーイングは、それが可能な場合は好ましい。位置キーイングは、CD-ROM（コンパクトディスク読み取り専用メモリ）などの記憶装置から出力記号の群を読み出すために十分機能し得る。ここで、出力記号の1群のキーは、CD-ROM上の位置（すなわち、トラック、プラスセクタ、セクタ内のプラス位置など）である。位置キーイングはまた、ATM（非同期転送モード）システムなどの回路ベース送信システムのために十分機能し得る。ここで、整列したデータセルは、厳しいタイミング制約下で送信される。この形態のキーイングを使用すると、受信者は、出力記号の1群のキーを、そのキーを明確に送信するのに必要なスペースを有することなく、再生成し得る。当然、位置キーイングは、そのような位置情報が利用でき確実にあることを必要とする。

20

**【0067】**

位置によるキーイングはまた、他のキーイング方法と組み合わせ得る。例えば、各パケットが出力記号の1より多くの群を含むパケット送信システムを考える。この場合、出力記号の群のキーは、固有のストリーム識別子、シーケンス番号、および出力記号の群のパケット内の位置から構築され得る。データイレーズは一般に全パケットの損失を生じるので、受信者は一般にパケットを全部受信する。この場合、受信者は、パケットのヘッダ（固有のストリーム識別子を含む）、パケット中のシーケンス番号、および出力記号の群のパケット内の位置を再生成し得る。

**【0068】**

いくつかのシステムにおいて好ましいキーイングの別の形態は、ランダムキーイングである。これらのシステムにおいては、乱数（疑似乱数）が、生成され、出力記号の各群に対するキーの少なくとも1部として使用され、および出力記号の群とともに明確に送信される。異なる物理的位置で異なる送信者によって生成されるキーに対してさえ（可能なキーの範囲が十分大きいと仮定する）、ランダムキーイングの1つの性質は、同じ値を有するキーの割合が小さい可能性がある。この形態のキーイングは、インプリメンテーションが簡単であるため、いくつかのシステムにおいては他の形態よりも利点を有し得る。

30

**【0069】**

上記のように、群連鎖反応符号化が有用であるのは、データイレーズの可能性があるか、または受信者が、送信が開始および終了するちょうどその時点で受信を開始および終了しない場合である。後者の条件は、本明細書中で「データ不完全性」と呼ばれる。これらの条件は、群連鎖反応符号化が使用される場合、通信プロセスに悪影響を与えない。なぜなら、受信される群連鎖反応符号化が、情報追加型であるように、極めて独立しているからである。出力記号の群のほとんどの任意集合が十分に独立しているので、大部分が情報追加型である（これは、本明細書中に記載の群連鎖反応符号化システムに対する場合である）ならば、任意の適切な数のパケットを使用して、入力ファイルを回復し得る。100パケットがデータイレーズを起こすノイズの発生によって損失する場合、その発生の後でさらに100パケットを取り上げてイレーズされたパケットの損失を置き換え得る。送信器が送信を開始した際に受信器が送信器に同調できず何千のパケットが損失した場合、受信器は、他の送信期間から（あるいは、別の送信器からの場合もある）その何千のパケットを取り上げる。群連鎖反応符号化を使用すると、受信者は、任意の特定のパケットのセッ

40

50

トを取り上げることに制約されず、従って1送信器からいくつかのパケット受信し、別の送信器に切り換え、いくつかのパケットを損失し、所定の送信の開始または終了を逸しても、なお入力ファイルを回復し得る。受信器 - 送信器調整のない送信に接続および切断する機能は、通信プロセスを大きく簡略化する。

#### 【0070】

(基本的な実施)

群連鎖反応符号化を使用してファイルを送信するステップは、入力ファイルから入力記号を生成、形成、または抽出し、その入力記号を1以上の出力記号の群に符号化するステップ(ここで、出力記号の各群は、出力記号のすべての他の群とは独立なキーに基づいて生成される)、および出力記号の群を1以上の受信者にチャネルを介して送信するステップを含む。群連鎖反応符号化を使用する入力ファイルの1コピーを受信(および再構築)するステップは、1以上のデータストリームから出力記号の群のあるセットまたはサブセットを受信するステップ、および受信された出力記号の群の値およびキーから入力記号を復号化するステップを含む。

10

#### 【0071】

以下に説明するように、デコーダは、出力記号の1以上の群の値および可能ならすでに回復された他の入力記号の値についての情報から入力記号を回復し得る。したがって、デコーダは、出力記号のいくつかの群からいくつかの入力記号を回復し得る。これにより、デコーダは、その復号された入力記号およびすでに受信された出力記号の群などからの他の入力記号を復号し得、したがってこれにより、ファイルの入力記号の回復の「連鎖反応」が受信者側で再構築される。

20

#### 【0072】

次に、本発明の局面を図面を参照して説明する。

#### 【0073】

図1は、連鎖反応符号化を使用する通信システム100のブロック図である。

通信システム100において、入力ファイル101、または入力ストリーム105は、入力記号発生器110に与えられる。入力記号発生器110は、入力ファイルまたは入力ファイルストリームから1以上の入力記号(IS(0), IS(1), IS(2), ...)のシーケンスを生成し、各入力記号は、値および位置を有する(図1に括弧付きの整数として示される)。上記のように、入力記号に対して可能な値、すなわち、そのアルファベットは、一般に $2^M$ 個の記号のアルファベットであるので、各入力記号は、Mビットの入力ファイルを符号化する。Mの値は、一般に通信システム100を使用することによって決定されるが、汎用システムは、入力記号発生器110に入力される記号サイズを含み得るので、Mは使用ごとに異なり得る。入力記号発生器110の出力は、エンコーダ115に与えられる。

30

#### 【0074】

キー発生器120は、エンコーダ115によって生成される出力記号の各群に対するキーを生成する。各キーは、どのキー発生器によって生成されたかにかかわらず、上記の方法のうちの1つ、または同じ入力ファイルに対して生成されたキーの大きな割合が固有であることを確保する任意の同等の方法によって生成される。例えば、キー発生器120は、カウンタ125の出力、固有のストリーム識別子130、および/またはランダム発生器135の出力の組合せを使用して、各キーを生成し得る。キー発生器120の出力は、エンコーダ115に与えられる。

40

#### 【0075】

キー発生器120によって与えられた各キーIに基づいて、エンコーダ115は、入力記号発生器によって与えられる入力記号から記号値B(I)のセットを有する1群の出力記号を生成する。出力記号の各群のb値は、そのキーおよび1以上の入力記号のある関数に基づいて生成される。本明細書中でb値を出力記号の群の「関連付けした入力記号」または単にその「関連付け」と呼ぶ。関数(「値関数」)および関連付けの選択は、以下により詳細に記載されるプロセスにしたがってなされる。一般に、常にではないが、Mは、入

50

力記号および出力記号に対して同じである、すなわち、入力記号は、出力記号と同じ長さである。

【 0 0 7 6 】

いくつかの実施形態において、入力記号数 K は、関連付けを選択するためにエンコードによって使用される。K があらかじめ未知の場合は（入力がストリーミングファイルの場合など）、K は単に推定され得る。値 K はまた、入力記号に対して記憶装置を割り当てるためにエンコード 1 1 5 によって使用され得る。

【 0 0 7 7 】

エンコード 1 1 5 は、出力記号の群を送信モジュール 1 4 0 に与える。送信モジュール 1 4 0 にはまた、キー発生器 1 2 0 から出力記号の各そのような群のキーが与えられる。送信モジュール 1 4 0 は、出力記号の群を送信し、送信モジュール 1 4 0 はまた、使用されるキーイング方法に依存して、出力記号の送信される群のキーについてのいくつかのデータをチャンネル 1 4 5 を介して受信モジュール 1 5 0 に送信する。チャンネル 1 4 5 は、イレーズチャンネルと仮定されるが、これは、通信システム 1 0 0 の適切な動作に必要なではない。モジュール 1 4 0、1 4 5 および 1 5 0 は、送信モジュール 1 4 0 が出力記号の群およびそれらのキーについて必要な任意のデータをチャンネル 1 4 5 に送信するように改造され、受信モジュール 1 5 0 が出力記号の群および可能ならばそれらのキーについてのいくつかのデータをチャンネル 1 4 5 から受信するように改造される限り、適切なハードウェアコンポーネント、ソフトウェアコンポーネント、物理的媒体、またはそれらの任意の組合せならいずれのものでもよい。K の値が、関連付けを決定するために使用される場合、チャンネル 1 4 5 を介して送信され得るか、またはエンコード 1 1 5 とデコード 1 5 5 との一致によってあらかじめ設定され得る。

【 0 0 7 8 】

上記のように、チャンネル 1 4 5 は、インターネット、またはテレビジョン送信器からテレビジョン受信者への放送リンク、または 1 地点から別の地点への電話接続などの実時間チャンネルであり得る。または、チャンネル 1 4 5 は、C D - R O M、ディスクドライブ、ウェブサイトなどの記憶チャンネルであり得る。チャンネル 1 4 5 は、実時間チャンネルと記憶チャンネルとの組合せでもあり得る。例えば、一人がパソコンからインターネットサービスプロバイダ（I S P）に電話回線を介して入力ファイルを送信し、その入力ファイルがウェブサーバに格納されて、後でインターネットを介して受信者に送信される場合に形成されるチャンネルである。

【 0 0 7 9 】

チャンネル 1 4 5 はイレーズチャンネルであると仮定されるので、通信システム 1 0 0 は、受信モジュール 1 5 0 を出る出力記号の群と送信モジュール 1 4 0 に入る出力記号の群との間に 1 対 1 対応を仮定しない。実際には、チャンネル 1 4 5 がパケットネットワークを含む場合、通信システム 1 0 0 は、任意の 2 つ以上のパケットの相対的順序がチャンネル 1 4 5 を介した送信中に保存されると仮定できないとされ得る。したがって、出力記号の群のキーは、上記キースキームの内 1 つ以上を使用して決定され、受信モジュール 1 5 0 を出る出力記号の群の順序によって必ずしも決定されない。

【 0 0 8 0 】

受信モジュール 1 5 0 は出力記号の群をデコード 1 5 5 に与え、受信モジュール 1 5 0 が受信する出力記号のこれらの群のキーについての任意のデータは、キー発生器 1 6 0 に与えられる。キー発生器 1 6 0 は、受信した出力記号の群に対するキーを再生成し、これらのキーをデコード 1 5 5 に与える。デコード 1 5 5 は、キー発生器 1 6 0 によって与えられたキーと対応の出力記号の群とを使用して、入力記号（再度、I S（0），I S（1），I S（2），...）を回復する。デコード 1 5 5 は、回復された入力記号を入力ファイルリアセンブラ 1 6 5 に与え、リアセンブラ 1 6 5 は入力ファイル 1 0 1 または入力ストリーム 1 0 5 のコピー 1 7 0 を生成する。

【 0 0 8 1 】

（基本エンコード）

図2は、図1に示されるエンコーダ115の1つの実施形態のブロック図である。図2のブロック図を、ここで図3を参照して説明する。図3は、図2に示すエンコーダによって実行される処理の内のいくつかの論理等価物を示す図である。

#### 【0082】

エンコーダ115には、入力記号および生成すべき出力記号の各群に対するキーが与えられる。示されるように、K個の入力記号が、入力記号バッファ205内に格納される。キーI（図1で示されるキー発生器120によって与えられる）は、値関数セクタ210、群サイズセクタ212、重みセクタ215、および連想装置220への入力である。入力記号数Kはまた、これら4つのコンポーネント210、212、215および220に与えられる。計算器225は、値関数セクタ210、群サイズセクタ212、重みセクタ215、連想装置220および入力記号バッファ205からの出力を受信するよう結合され、出力記号値の群に対する出力を有する。理解すべきことは、図2に示す要素と等価な他の構成が使用され得、これは本発明のエンコーダの1例にすぎないことである。

10

#### 【0083】

動作中、K個の入力記号が入力記号バッファ205に受信および格納される。上記のように、各入力記号は、位置（すなわち、入力ファイル内の元の位置）および値を有する。入力記号は、格納される入力記号の位置が決定され得る限り、それぞれの順序で入力記号バッファ205内に格納される必要はない。

#### 【0084】

20

群サイズセクタ212は、使用される場合、キーIおよび入力記号数Kから群サイズG(I)を決定する。1つの変形において、群サイズG(I)は、すべてのIに対して同じである（すなわち、 $G(I) = b$ であり、ここでbは正の整数である）。この変形において、群サイズセクタ212は必要ではなく、計算器225はbの値に設定され得る。例えば、群サイズは、すべてのIに対して4に設定され得る、すなわち、 $G(I) = b = 4$ であり、出力記号の各群は、4つの出力記号を含む。いくつかの実施形態において、2または3などの小さな値のb、またはわずかに異なる値のG(I)を使用することは、エンコーダの効率化に役立つ。b = 1の場合、その結果は、 $Luby\ I$ に記載のエンコーダと実質的に同じである。

#### 【0085】

30

キーIおよび入力記号数Kを使用して、重みセクタ215は、キーIを有する出力記号の群の「関連付け」となるべき入力記号の数 $W(I)$ を判定する。キーI、重み $W(I)$ および入力記号数Kを使用して、連想装置220は、出力記号の群に関連付けられた入力記号の位置のリスト $AL(I)$ を判定する。理解すべきことは、 $W(I)$ は、連想装置220があらかじめ $W(I)$ を知らずに $AL(I)$ を生成し得る場合、別個にまたは明確に計算される必要がないことである。 $AL(I)$ が一旦生成されると、 $W(I)$ は容易に判定される。なぜなら、それは $AL(I)$ 中の関連付けの数であるからである。

#### 【0086】

I、 $W(I)$ および $AL(I)$ が一旦既知となると、出力記号の群のG(I)個の値 $B(I) = B_{-1}(I), \dots, B_{-G(I)}(I)$ は、値関数F(I)210に基づいて計算器225によって計算される。適切な値関数の1つの性質は、 $AL(I)$ 中のG(I)個までの関連付けに対する未知の値が出力記号値 $B(I)$ の群および $AL(I)$ 中の他の関連付けに対する既知の値から判定され得ることである。このステップにおいて使用される1つの好ましい値関数は、リード・ソロモン値関数である。なぜなら、それは、この性質を満足し、容易に計算され、および容易に反転されるからである。

40

#### 【0087】

リード・ソロモン値関数の場合、値 $B(I)$ は、リード・ソロモン符号化スキームにしたがって、 $AL(I)$ 中の関連付けの値から計算される。そのスキームにおいて、リード・ソロモン冗長記号は $B(I)$ であり、そして $AL(I)$ 中の入力記号とともにこれらの記号は、リード・ソロモン符号化スキームを使用して符号化される符号化ブロックを形成す

50

る。

#### 【0088】

他の適切な値関数を代わりに使用してもよい。例えば、群サイズが1に等しい場合のXOR値関数の使用、有限体に関する多項式に基づく方法、線形系方程式に基づく方法、有限体に関するコーシー行列に基づく方法、または、他のMDS符号（そのリード・ソロモン符号は1例である）を含む。

#### 【0089】

出力記号の群の関連付け数に依存する異なる値関数の混合がまた、使用され得る。例えば、関連付け数 $W(I)$ が $G(I)$ に等しい場合、識別関数を使用され得る、すなわち、 $i = 1, \dots, G(I)$ に対して、その群における $i$ 番目の出力記号の値は $B\_i(I) = IS(P\_i)$ であり、ここで $P\_1, \dots, P\_G(I)$ は $G(I)$ 個の関連付けの位置である。関連付け数 $W(I)$ が $G(I) + 1$ に等しい場合、 $i = 1, \dots, G(I)$ に対して、その群における $i$ 番目の出力記号の値は、 $B\_i(I) = IS(P\_i) XOR IS(P\_i + 1)$ であり得る、ここで、 $P\_1, \dots, P\_G(I) + 1$ は、 $G(I) + 1$ 個の関連付けの位置である。関連付け数 $W(I)$ は $G(I) + 1$ よりも大きい場合、リード・ソロモン値関数を使用され得る。

#### 【0090】

値関数セクタ210が、使用される場合、キー $I$ および $K$ からの値方法または値関数 $F(I)$ を判定し、ここで、 $F(I)$ を使用して $I$ に対する値を計算する。1つの変形において、値関数セクタ210は、すべての $I$ に対して同じ方法または関数 $F$ を使用する。この変形において、値関数セクタ210は必要でなく、計算器225が値関数 $F$ を構成し得る。例えば、値関数はすべての $I$ に対してリード・ソロモン符号化器であり得る。すなわち、出力記号値の群は、すべてのその関連付けの値に基づいてリード・ソロモン符号化器によって計算される $G(I)$ 個の冗長記号を含む。

#### 【0091】

各キー $I$ に対して、重みセクタ215は、 $I$ および $K$ から重み $W(I)$ を判定する。1つの変形において、重みセクタ215は、キー $I$ を使用することによって $W(I)$ を選択して、まずランダム検索数を生成し、次いでこの数を使用して、重みセクタ215内に格納された分布テーブル中の $W(I)$ の値を検索する。そのような分布テーブルがどのように形成およびアクセスされるかを以下により詳細に説明する。一旦重みセクタ215が $W(I)$ を判定すると、この値は、連想装置220および計算器225に与えられる。

#### 【0092】

連想装置220は、現在の出力記号に関連付けられた $W(I)$ 個の入力記号の位置のリスト $AL(I)$ を判定する。この関連付けは、 $I$ の値、 $W(I)$ の値および $K$ （利用可能ならば）に基づく。一旦連想装置220が $AL(I)$ を判定すると、 $AL(I)$ は、計算器225に与えられる。リスト $AL(I)$ 、重み $W(I)$ および、値関数セクタ210によって与えられる値関数 $F(I)$ または予め選択された値関数 $F$ のいずれかを使用して、計算器225は、入力記号バッファ205中の $AL(I)$ によって参照される $W(I)$ 個の入力記号にアクセスして、出力記号の現在の群に対する値 $B(I) = B\_1(I), \dots, B\_b(I)$ を計算する。ここで $b$ は、群サイズセクタ212によって与えられる群サイズ $G(I)$ または予め選択された設定値のいずれかである。 $AL(I)$ を計算するための手順の例を以下に与えるが、別の適切な手順を代わりに使用してもよい。好ましくは、その手順は、各入力記号に、出力記号の所定の群に対する関連付けとして選択される確率を略等しく与え、デコーダがそれにとって利用可能な $AL(I)$ をまだ有していない場合にデコーダが複製し得るように選択する。

#### 【0093】

次に、エンコーダ115は、 $B(I) = B\_1(I), \dots, B\_b(I)$ を出力する。実際は、エンコーダ115は、図3に例示する動作を実行して、すなわち、選択された入力記号のある値関数として出力記号値の群、 $B(I) = B\_1(I), \dots, B\_b$

10

20

30

40

50

(I)を生成する。示される例において、その値関数は、群サイズbが2に設定されたリード・ソロモン符号化であり、出力記号の群の重みW(I)は3であり、関連付けられた入力記号(関連付け)は位置0、2および3であり、値IS(0)、IS(2)およびIS(3)をそれぞれ有する。したがって、出力記号の群は、Iの値に対して以下のように計算される：

$$B\_1(I) = RS\_1(IS(0), IS(2), IS(3))$$

$$B\_2(I) = RS\_2(IS(0), IS(2), IS(3))$$

ここで、RS\_i(\*)は、リード・ソロモンエンコーダを引数に適用してi番目の冗長記号を生成することによって生成される記号値である。

【0094】

次に、生成された出力記号の群は、上記のように送信および受信される。ここで、出力記号の群の内いくつかが損失されるかまたはその順序がみだれた、もしくは1つ以上のエンコーダによって生成されたと仮定する。しかし、受信される出力記号の群の受信は、そのキーIの表示、および値B(I) = B\_1(I), ..., B\_b(I)が正確であるという何らかの確実さをもって受信されると仮定する。図1に示すように、それらの出力記号の受信された群は、キー再発生器160による表示およびKの値から再構築される対応のキーとともに、デコーダ155に入力される。

【0095】

入力記号中に符号化されるビット数M(すなわち、そのサイズ)は、アプリケーションに依存する。出力記号のサイズおよび1つの群中の出力記号数bもまた、アプリケーションに依存し得るが、チャンネルにもまた依存し得る。例えば、一般的な入力ファイルが、複数のメガバイトのファイルである場合、その入力ファイルは、数千個、数万個または数十万個の入力記号に分割され、各入力記号は、数バイト、数十バイト、数百バイトまたは数千バイトである。

【0096】

いくつかの場合において、出力記号値および入力記号値が同じサイズであれば(すなわち、同じ数のビットによって表現可能か、または同じアルファベットから選択される)、符号化プロセスは、簡略化され得る。そのような場合、出力記号の群をパケットに入れるのが望ましく、かつ出力記号の各群が限られたサイズのパケットに収まる必要があり、かつ各群のサイズbが固定値に設定される場合などの出力記号値サイズが限定される場合、入力記号値サイズが限定される。キーについてのいくつかのデータが、受信器にてキーを回復するために送信される場合、好ましくは、出力記号の群は、値およびキーについてのデータが1つのパケットに収まるように十分に小さいものであり得る。

【0097】

上記のように、多くの実施において入力記号の位置は一般に連続であるが、キーは連続ではない。例えば、入力ファイルが60,000個の入力記号にまで分割されたとすると、その入力記号の位置の範囲は、0から59,999ある。上記実施のうちの1つにおいて、各キーはランダムな32ビット数として独立に選択され、出力記号の群は、連続的に生成され、送信者が停止するまで送信され得る。ここで示すように、群連鎖反応符号化は、60,000記号の入力ファイルが出力記号の任意の十分に大きな集合(60,000 + いくつかの増分A)から再構築される。なお、これは、出力記号のこれらの群が出力シーケンスのどこから取り出されるかに依存しない。

【0098】

(基本デコーダ)

図4は、デコーダ155の1つの実施形態を詳細に示し、多くの部分が図2に示すエンコーダ115と共通である。デコーダ155は、値関数セクタ210、群セクタ212、重みセクタ215、連想装置220、出力記号の1群を格納するバッファ405、レジュース(reducer)415、再構築器420および再構築バッファ425を含む。エンコーダに関して、群セクタ212および群サイズを格納するために割当てられるバッファ405中のスペースは、群のサイズが出力記号のすべての群に対して同じである

10

20

30

40

50

場合、任意であり、使用しなくてもよい。同様に、値関数セクタ 2 1 0 および値関数の記述を格納するために割当てられるバッファ 4 0 5 中のスペースは、値関数が出力記号のすべての群に対して同じである場合、任意であり、使用しなくてもよい。再構築バッファ 4 2 5 のいくつかのエントリが図示され、ここでいくつかの入力記号が再構築され、他は依然未知である。例えば、図 4 において、位置 0、2、5 および 6 の入力記号が回復され、位置 1、3 および 4 の入力記号はまだ回復されない。

#### 【 0 0 9 9 】

動作中、キー  $I$  および値  $B(I) = B_{-1}(I), \dots, B_{-b}(I)$  を有する受信された出力記号の各群に対して、デコーダ 1 5 5 は以下を行う。キー  $I$  は、値関数セクタ 2 1 0、群セクタ 2 1 2、重みセクタ 2 1 5 および連想装置 2 2 0 に与えられる。K およびキー  $I$  を使用して、重みセクタ 2 1 5 は、重み  $W(I)$  を判定する。K、キー  $I$  および重み  $W(I)$  を使用して、連想装置 2 2 0 は、出力記号の群に関連付けられた入力記号の  $W(I)$  個の位置のリスト  $AL(I)$  を生成する。必要に応じて、K および  $I$  を使用して、群セクタ 2 1 2 は、群サイズ  $G(I)$  を選択する。必要に応じて、K および  $I$  を使用して、値関数セクタ 2 1 0 は、値関数  $F(I)$  を選択する。次に、 $I$ 、 $B(I)$ 、 $W(I)$  および  $AL(I)$ 、任意の  $G(I)$  および任意の  $F(I)$  は、列をなしてバッファ 4 0 5 に格納される。群セクタ 2 1 2、値関数セクタ 2 1 0、重みセクタ 2 1 5 および連想装置 2 2 0 は、エンコーダ 1 1 5 について記載したのと同じようにデコーダ 1 5 5 の動作を実行する。特に、図 5 における群セクタ 2 1 2、値関数セクタ 2 1 0、重みセクタ 2 1 5 および連想装置 2 2 0 によって生成される群サイズ  $G(I)$ 、値関数  $F(I)$ 、重み  $W(I)$  およびリスト  $AL(I)$  は、図 4 に示す対応の部分と同じキー  $I$  である。K が入力ファイルごとに異なる場合、K をメッセージヘッダに含むなどの任意の従来方法でエンコーダからデコーダへ通信され得る。

#### 【 0 1 0 0 】

再構築器 4 2 0 は、バッファ 4 0 5 を走査して、たかだか群サイズの重みを有するバッファ 4 0 5 に格納された出力記号の群を探す。これらの記号は、本明細書中で「復号可能なセット」のメンバと呼ぶ。上記の性質を有する値関数に対して、たかだかそれらの群サイズの重みの出力記号の群は、復号可能なセットに含まれる。なぜなら、それらの群サイズまでの入力記号の未知の値は、その出力記号の群および関連付けである他の入力記号の既知の値から判定され得るからである。当然、入力記号が、たかだかそれらの群サイズの重みを有すること以外の条件下で出力記号の 1 群によって復号されることを可能にし得る値関数が使用されるならば、その条件を使用して、出力記号の 1 群が復号可能なセットかどうかを判定し得る。明確にするために、ここで記載した例は、復号可能なセットが、たかだかそれらの群サイズの重みを有する出力記号の群であり、これらの例の、他の値関数復号可能条件への拡張は、この記載から明らかである。

#### 【 0 1 0 1 】

再構築器 4 2 0 が復号可能なセットに含まれるキー  $I$  を有する出力記号の 1 群を見つけた場合、その出力記号値の群は、 $B(I) = B_{-1}(I), \dots, B_{-G}(I)$  であり、必要に応じて、値関数  $F(I)$  を使用して  $AL(I)$  にリストされた未知の入力記号の  $W(I)$  個の値を再構築し、再構築された入力記号は、それらの入力記号に対する適切な位置で再構築バッファ 4 2 5 に配置される。 $W(I)$  が厳密に群サイズ  $G(I)$  より小さい場合、再構築器 4 2 0 は、独立した方法ですでに再構築された入力記号値を再構築し得る。この場合、再構築器 4 2 0 は、独立に再構築された入力記号を脱落させ、既存の再構築された入力記号を上書きし、または独立に得られた入力記号の値を比較し、それらが異なれば誤りを生成する。このように再構築器 4 2 0 は入力記号を再構築するが、復号可能なセット中の出力記号の群からのみ入力記号を再構築する。一旦復号可能なセットからの出力記号の 1 群を使用して入力記号を再構築すると、それは、バッファ 4 0 5 のスペースを節約するために削除され得る。「使い古した」群の出力記号を削除することはまた、再構築器 4 2 0 が連続してその出力記号の群に再び訪れないことを確実にする。

#### 【 0 1 0 2 】

まず、再構築器 4 2 0 は、復号可能なセットの 1 メンバである少なくとも 1 つの出力記号の群が受信されるまで待機する。一旦その出力記号の 1 群を使用して未知の入力記号を再構築すると、復号可能なセットは再び空となる。ただし、出力記号のいくつかの他の群が、これらのちょうど再構築された入力記号の内いくつかおよびその群サイズまでまだ再構築されない他の入力記号の関数であり得るという事実を除く。このように、復号可能なセットの 1 メンバから入力記号を再構築することにより、出力記号の他の群が復号可能なセットに付加される。出力記号の群を復号可能なセットに付加して低減するプロセスは、レジューサ 4 1 5 によって実行される。

#### 【 0 1 0 3 】

レジューサ 4 1 5 は、バッファ 4 0 5 および再構築バッファ 4 2 5 を走査して、回復された入力記号の位置をリストするリスト A L ( I ) を有する出力記号の群を見つける。レジューサ 4 1 5 がキー I を有する出力記号のそのような「低減可能な」群を見つけた場合、レジューサ 4 1 5 は、A L ( I ) 中の入力記号のうちでまだ再構築されていなかった入力記号値の数に等しくなるように W ( I ) を再計算する。

#### 【 0 1 0 4 】

レジューサ 4 1 5 の作用は、バッファ 4 0 5 中の出力記号の群の重みを低減する。出力記号の 1 群の重みは、たかだかその群サイズに低減される場合（または、他の復号可能な条件が他の値関数に対して生じる）、その出力記号の群は、復号可能なセットの 1 メンバとなり、それは次に再構築器 4 2 0 によって作用され得る。実用においては、一旦十分な数の出力記号の群が受信されると、レジューサ 4 1 5 および再構築器 4 2 0 は、連鎖反応復号化を生成し、入力ファイルからのすべての入力記号が回復されるまで、再構築器 4 2 0 は復号可能なセットを復号してより多くの入力記号を回復し、レジューサ 4 1 5 はそれらの新しく受信された入力記号を使用してより多くの出力記号の群を低減し、それによりそれらを復号可能なセットに付加するなどする。

#### 【 0 1 0 5 】

図 4 に示されるデコーダは、記憶装置、計算サイクルまたは送信時間をあまり考慮せずに、単純な方法で入力記号を再構築する。デコーダメモリ、複合化時間または送信時間（受信される出力記号の群の数を制限する）が限定される場合、デコーダは、それら限られたリソースをよりよく使用するために最適化され得る。

#### 【 0 1 0 6 】

（より効率的なデコーダ）

図 5 は、デコーダ 5 0 0 のより効率的な実施の好ましい実施形態を詳細に示す。ここで、値関数は、リード - ソロモン値関数であると仮定する。さらに効率的になる可能性のある同様の実施が、リード - ソロモン値関数以外の値関数に対して適用される。図 5 を参照して、デコーダ 5 0 0 は、出力記号の群データ構造 5 0 5（以下、G O S D S 5 0 5 と呼ぶ）、入力記号データ構造 5 1 0（以下、I S D S 5 1 0 と呼ぶ）、復号可能なセットスタック 5 1 5（以下、D S S 5 1 5 と呼ぶ）、受信オーガナイザ 5 2 0、および回復プロセス 5 2 5 を含む。

#### 【 0 1 0 7 】

G O S D S 5 0 5 は、出力記号の群についての情報を格納するテーブルであり、ここで G O S D S 5 0 5 の行 R は、受信された出力記号の R 番目の群についての情報を格納する。変数 R は、受信された出力記号の群の数を追跡し、ゼロに初期化される。G O S D S 5 0 5 は、各行に対してフィールド K E Y、V A L U E、および W E I G H T を格納し、図示のフィールドは、列に組織化される。K E Y フィールドは、出力記号の群のキーを格納する。V A L U E フィールドは、出力記号値の群を格納する。W E I G H T フィールドは、出力記号の群の初期重みを格納する。出力記号の 1 群の W E I G H T は、たかだか群サイズになるまで経時的に低減され、そして入力記号を回復するために使用され得る。

#### 【 0 1 0 8 】

I S D S 5 1 0 は、入力記号について情報を格納するテーブルであり、ここで行 P は、位置 P での入力記号についての情報を格納する。各行に対して、I S D S 5 1 0 は、最終的

10

20

30

40

50

に回復された入力記号の値になる  $REC\_VAL$  フィールド、すべての値が「no」に初期化され、入力記号が回復されたかどうかを示す  $REC\_IND$  フィールド、および  $RL$  フィールド用の記憶装置を含む。入力記号が回復される場合、入力記号の  $REC\_IND$  は、「yes」に変更される。 $RL$  列は、すべてが「空リスト」値に初期化される。関連付けとして入力記号を有する出力記号の群が受信されるので、出力記号の群の  $GOSDS505$  における行数が、入力記号に対する  $RL$  リストに付加される。

#### 【0109】

$DS515$  は、復号可能なセットについての情報を格納するスタックである。変数  $S$  は、復号可能なセットのサイズを追跡し、それをゼロに初期化する。 $DS515$  において、列  $OUT\_ROW$  は、出力記号の群の  $GOSDS505$  における行数を格納する。

10

#### 【0110】

1つの実施形態において、デコーダ500は、以下のように、そして図6のフローチャートに示すように動作され、図6の対応するステップは、プロセスの説明において括弧で示す。まず、 $ISDS510$  は、上記のように初期化され、 $R$  および  $S$  の両方がゼロに初期化される(605)。出力記号の1つの新しい群が受信されると(610)、すなわち、キー  $I$  および出力記号値の群  $B(I) = B\_1(I), \dots, B\_G(I)(I)$ 、 $GOSDS505$  において  $KEY(R)$  が  $I$  に設定され、 $VALUE(R)$  が  $B(I) = B\_1(I), \dots, B\_G(I)(I)$  に設定される(615)。次に、受信オーガナイザ520が呼び出されて  $GOSDS505$  の行  $R$  に格納されたキー  $I$  を有する受信された出力記号の群が処理される(620)。これは、図7のフローチャートに示すように、 $ISDS510$  に格納された情報を適切に使用して、情報を  $GOSDS505$  および  $DS515$  に付加するステップを含む。次に、 $R$  を1だけインクリメントし(625)、次の出力記号の群を  $GOSDS505$  の次の行に格納させる。次に、回復プロセッサ525が呼び出され、復号可能なセット中の出力記号の群を処理し、出力記号の新しい群を復号可能なセットに付加する(630)。これは、図8aおよび/または8bのフローチャートに示すように、 $ISDS510$  および  $GOSDS505$  の一部を適切に使用および改変することによって、 $DS515$  に格納された復号可能なセットに付加および削除するステップを含む。デコーダ500は、回復された入力記号の数を追跡し、この数が  $K$  に達した場合、すなわち、すべての入力記号が回復された場合、デコーダ500は、首尾よく終了し、そうでなければステップ610に戻り次の出力記号の群を受信する(635および640に示す)。

20

30

#### 【0111】

受信オーガナイザ520の動作を説明するフローチャートを図7に示す(図9から12を参照)。値  $B(I) = B\_1(I), \dots, B\_G(I)(I)$  およびキー  $I$  を有する出力記号の1群が到着した場合、受信オーガナイザ520は、以下の動作を実行する(図7を参照)。重み  $W(I)$  は、 $I$  および  $K$  から計算され(705)、関連付けの位置のリスト  $AL(I)$  は、 $I$ 、 $W(I)$ 、および  $K$  から計算される(710)。図11~12は、 $W(I)$  の1計算の詳細を示し、図9~10は、 $AL(I)$  の1計算の詳細を示す。

#### 【0112】

再度図7を参照して、リスト  $AL(I)$  上の各位置  $P$  に対して、入力記号  $P$  が回復されない場合、すなわち、 $ISDS510$  において  $REC\_IND(P) = \text{「no」}$  の場合、 $R$  は、 $ISDS510$  中のリスト  $RL(P)$  に付加され、そうでなく、入力記号  $P$  が回復される場合、すなわち、 $ISDS510$  において  $REC\_IND(P) = \text{「yes」}$  の場合、 $W(I)$  は1だけデクリメントされる(720)。次に、 $WEIGHT(R)$  は、 $GOSDS505$  において  $W(I)$  に設定される(725)。次に、 $WEIGHT(R)$  は、 $G(I)$  と比較される(730)。 $WEIGHT(R)$  がたかだか  $G(I)$  である場合、出力記号の群は、復号可能なセットに付加される、すなわち、 $OUT\_ROW(S)$  が  $DS515$  において  $R$  に設定され、 $S$  の値が1だけインクリメントされる(735)。最後に、受信オーガナイザ520が戻る(740)。

40

#### 【0113】

50

図 8 a に、図 9 ~ 12 を参照して回復プロセッサ 525 の 1 つの動作を記載するフローチャートを示す。その動作中、回復プロセッサ 525 はまず、復号可能なセットが空かどうか、すなわち、 $S = 0$  かどうかを確認するためにチェックし、そうであればただちに呼び出し元へ帰る (805、810)。復号可能なセットが空でなければ、 $S$  は 1 だけデクリメントされ (815)、出力記号の群の行番号  $R'$  を  $DS515$  からロードする (820)。次に、出力記号の元のキー、 $GOSDS505$  からの  $KEY(R')$  が  $I$  へロードされ (835)、元の重み  $W(I)$  および群の関連付けの元のリスト  $AL(I)$  を計算する (840、845)。  $AL(I)$  におけるのすべての入力記号がすでに回復された場合、すなわち、 $AL(I)$  におけるすべての  $P$  に対して  $REC\_IND(P) = 'yes'$  ならば (847)、回復プロセッサ 525 は、復号可能なセットのその要素の処理を停止し、次のステップを処理し続ける。そうでなければ、 $GOSDS505$  における行番号  $R'$  に格納された出力記号の群を使用して、 $ISDS510$  における  $AL(I)$  にある入力記号すべての未知の値を回復する。これは、リード-ソロモンデコードを利用することによって実行され、出力記号値の群および  $AL(I)$  における入力記号の既知の値を使用して  $AL(I)$  における  $G(I)$  個までの入力記号の未知の値を回復する (850)。次に、入力記号値が回復されるすべての位置  $P$  に対して、 $REC\_VAL(P)$  は、 $ISDS510$  における回復された入力記号値に設定され (852)、 $REC\_IND(P)$  が 'yes' に設定され、入力記号が  $ISDS510$  において回復されたことを示す (852)。

#### 【0114】

図 8 a において示すプロセスの 1 つの変形例を図 8 b に示す。そこで、出力記号の各群が処理されるステップ 850、852、855 および 860 を実行する代わりに、図 8 b のステップ 851、853、856 および 861 に示すように、 $R'$  の値が後の処理のために実行スケジュールに格納される。保留された実行処理の例を、ステップ 870、875、880 および 885 により示されるステップを含む図 8 c に示す。この変形例において、図 6 に示すフローチャートは、ステップ 605 において  $E$  をゼロに初期化することによって改変される。実行スケジュールの保留された処理は、受信した記号がファイル全体を復号するのに十分であるとデコードが判定した後、例えば、すべての入力記号が回復可能であると分かった後のステップ 640 にて行われ得る。いくつかの場合において、特に入力記号が大きい場合、スケジュールの実行は、入力ファイルまたはその部分が受信器側で必要となるまで保留され得る。

#### 【0115】

いずれの変形例においても、すなわち、図 8 a または図 8 b のいずれかにおいて、ステップ 855 または 856 で、ちょうど回復された出力記号の群または回復されるはずだった出力記号の群に対し、関連付けとしての入力記号は、これらの入力記号が回復されたこと、または回復されるはずだったことを反映するように改変される。ちょうど回復された、または回復されるはずだった各位置  $P$  に対して、 $GOSDS505$  における出力記号のこれらの群の行番号は、 $RL(P)$  内に格納される。 $RL(P)$  における各行番号  $R''$  に対して、 $WEIGHT(R'')$  が 1 だけデクリメントして、 $GOSDS505$  の行  $R''$  における出力記号の群の関連付けとして位置  $P$  における入力記号の回復を反映させる。この改変により、 $GOSDS505$  の行  $R''$  における出力記号の群が群サイズに等しい重み、すなわち、 $WEIGHT(R'') = G(KEY(R''))$  になる場合、出力記号のこの群は、 $OUT\_ROW(S)$  を  $R''$  に設定し、かつ  $S$  を 1 だけインクリメントすることによって復号可能なセットに付加される (855 または 856)。最後に、リスト  $RL(P)$  上の出力記号の群の行番号を格納するために使用されるスペースをフリースペースに戻し (860 または 861)、処理はステップ 805 に続く。

#### 【0116】

##### (連想装置の実施)

1 群の出力記号キーから関連する入力記号へのマッピング (すなわち、重み  $W(I)$ 、およびキー  $I$  に対する関連付けの位置のリスト  $AL(I)$  の決定) は、種々の形態をとり得

る。 $W(I)$ は、同じキー $I$ に対してエンコーダおよびデコーダの両方によって同じ値（送信者および受信者のそれぞれにおいて）であるように選択されるべきである。同様に、 $AL(I)$ は、同じキー $I$ に対してエンコーダおよびデコーダの両方によって同じ位置のリストを含むように選択されるべきである。連想装置は、 $I$ ならびに通常 $W(I)$ および $K$ から $AL(I)$ を計算または生成するオブジェクトである。

#### 【0117】

1つの実施形態において、 $W(I)$ および $AL(I)$ は、ランダムプロセスを模倣するように設計された方法で判定される。エンコーダおよびデコーダが同じキーに対して同じ結果を生成するという要件を満足するために、擬似ランダムシーケンスが、キーをシードにしてエンコーダおよびデコーダの両方によって生成され得る。擬似ランダムシーケンスの代わりに、本当のランダムシーケンスが、 $W(I)$ および/または $AL(I)$ の生成のために使用され得るが、それが有用であるためには、 $W(I)$ および $AL(I)$ を生成するために使用されたランダムシーケンスが受信者に通信される必要があり得る。

#### 【0118】

図4において示すデコーダにおいて、バッファ405は、関連付けの位置の出力記号リストの各群の記憶装置、すなわち列でラベルされた $AL(I)$ 内に記憶装置を必要とする。図5に示すより効率的なデコーダは、この記憶装置を必要としない。なぜなら、関連付けの1リストは、例えば図9～10に示すように、必要に応じて再計算されるからである。これらの計算が必要に応じて速やかに行われ得る場合にのみ、記憶装置を節約するために毎回関連付けリストを再計算する際に利点がある。

#### 【0119】

連想装置220の好ましい実施形態が図9に示され、図10に示すプロセスにしたがって動作する。この連想装置は、エンコーダおよびデコーダにおいて使用され得る。エンコーダは1回に1より多くの $AL(I)$ を格納する必要が通常ないので、エンコーダでの $AL(I)$ のための記憶装置にはあまり関心がないが、同じプロセスをエンコーダおよびデコーダの両方で使用して、 $AL(I)$ に対する値が両方の場所で確実に同じとなるようにすべきである。

#### 【0120】

連想装置220への入力は、キー $I$ 、入力記号数 $K$ 、および重み $W(I)$ である。出力は、キー $I$ を有する出力記号の群の関連付けの $W(I)$ 個の位置のリスト $AL(I)$ である。図9において示すように、アソシエータは、ランダムビットのテーブルASSOC\_\_RBITS905および計算器ASSOC\_\_CALC910を含む。特定の $AL(I)$ が生成される前に、入力ファイルのサイズは、入力記号数が素数となるように調節される。このように、入力ファイルが $K$ 個の入力記号で開始する場合、 $K$ より大きいまたは $K$ に等しい最小素数 $P$ が識別される。 $P$ が $K$ よりも大きい場合、 $P-K$ 個のブランク（例えば、ゼロに設定される）入力記号が入力ファイルに付加され、 $K$ が $P$ に再設定される。この変更された入力ファイルに対して、関連付けの位置のリスト $AL(I)$ は、図9および10に示すように計算される。

#### 【0121】

この実施形態において、ASSOC\_\_CALC910は、以下に記載され、かつ図10のフローチャートに示すように動作する。第1ステップは、キー $I$ 、入力記号数 $K$ およびランダムビットASSOC\_\_RBITS905のテーブルを使用して、 $X$ が少なくとも1かつたかだか $K-1$ であり、 $Y$ が少なくともゼロかつたかだか $K-1$ である性質を有する2つの整数値 $X$ および $Y$ を生成することである（1005）。好ましくは、 $X$ および $Y$ は、独立であり、それぞれの範囲内で一様分布する。次に、 $W(I)$ 個のエントリを有する配列 $V[]$ が、 $AL(I)$ の記憶のために、そのメンバが計算される際に初期化される（1010）。 $V[]$ は、1つのリストに対する一時記憶装置にすぎないので、バッファ405（図4を参照）の $AL(I)$ 列よりもずっと少ないメモリを占めるだけであり得る。 $V[0]$ （これは、リスト $AL(I)$ の第1番目の要素）が $Y$ に設定される（1015）。次に、1で始まり $W(I)-1$ で終わる $J$ のすべての値に対して、 $V[J]$ の値は、ステ

ップ1020～1050に示すように、 $(V[J-1] + X) \bmod K$ に設定される。Kが素数であり、かつ $W(I)$ がたかだかKであるので、 $V[]$ 値のすべてが固有である。図示のように、「 $\bmod K$ 」演算は、簡単な比較および減算演算であり得る、すなわち、ステップ1035および1040。このように、出力記号の所定群の関連付けの位置のリストを生成するプロセスは、非常に効率的である。

#### 【0122】

$AL(I)$ を計算する上記アプローチの1つの利点は、関連付けの位置の分布に十分な変化を生成して、復号化アルゴリズムを、 $W(I)$ を選択するための良好な手順と合わされる場合、高い確率かつ最小の受信オーバーヘッドで動作させることを確実にする（すなわち、入力ファイルは、出力記号の $K/b$ 個よりわずかに多い群を受信した後、回復される。ここで入力記号および出力記号は同じ長さとする）。

10

#### 【0123】

$AL(I)$ が図10に示すように計算され、ランダムビットASSOC\_\_RBITS905のテーブルが送信者と受信者との間で秘匿される場合、通信システム100は、安全な通信システムとして使用され得る。なぜなら、 $AL(I)$ のシーケンスを知らずに、受信したストリームを復号化することは、不可能ではないが、困難であるからである。この特徴はまた、 $AL(I)$ を計算または生成する方法である限り、 $AL(I)$ を計算または生成する他の方法を使用する通信システムに適用し得る、また、好ましくは、発生器 $AL(I)$ を計算する方法において使用されるシードは、秘匿される。

#### 【0124】

20

（重みセクタ実施例）

エンコーダ/デコーダの性能および効率は、重みの分布に依存し、いくつかの分布は、他の分布より良好である。重み選択の動作の局面を以下に議論し、その後いくつかの重要な重み分布を説明する。図11のブロック図および図12のフローチャートを使用してこれらの概念を例示する。

#### 【0125】

図11に示すように、重みセクタは、2つのプロセスWT\_\_INIT1105およびWT\_\_CALC1110、ならびに2つのテーブルWT\_\_RBITS1115およびWT\_\_DISTRIBUT1120を含む。プロセスWT\_\_INIT1105は、第1のキーが通過してテーブルWT\_\_DISTRIBUT1120を初期化した場合に1度だけ呼び出される。WT\_\_DISTRIBUT1120の設計は、システムの重要な局面であり、後でずっとより詳細に考察される。プロセスWT\_\_CALC1110は、各呼び出しごとに呼び出されて、キーIに基づいて重み $W(I)$ を生成する。図12のフローチャートに示すように、WT\_\_CALC1110は、キーおよびテーブルWT\_\_RBITSに格納されたランダムビットを使用して、乱数Rを生成する（1205）。次に、Rの値を使用して、テーブルWT\_\_DISTRIBUT1120における行番号Lを選択する。

30

#### 【0126】

図11に示すように、WT\_\_DISTRIBUT1120のRANGE列におけるエントリは、値MAX\_\_VALで終了する正の整数の増加シーケンスである。Rに対する可能な値のセットは、ゼロおよびMAX\_\_VAL-1との間の整数である。所望の性質は、Rが、可能な値の範囲においてどの値になるのも等しく確からしいことである。Lの値は、RANGE(L-1)  $R$  < RANGE(L)を満足するLを見つけるまでRANGE列を検索することによって決定される（1210）。一旦Lが見つかり、 $W(I)$ の値がWT(L)に設定され、これは戻された重みである（1215、1220）。図11において、図示のテーブル例に対して、Rが38,500に等しい場合、Lが4であることが分かり、したがって、 $W(I)$ は $WT(4) = 8$ に設定される。

40

#### 【0127】

重みセクタおよび連想装置を実施する他の変形例は、Iを擬似ランダムに生成し、Iから $W(I)$ および $AL(I)$ を直接生成するステップを含む。明らかなように、 $W(I)$ は、 $AL(I)$ を調べることによって決定され得る。なぜなら、 $W(I)$ は、 $AL(I)$

50

における関連付けの数に等しいからである。この説明から、 $W(I)$  個の値を生成する多くの他の方法は、説明したばかりのシステムと等価であり、 $WT\_DISTRIB$ によって規定される分布を有する 1 セットの  $W(I)$  値を生成するということが明らかである。

【0128】

$W(I)$  が図 12 に示すように計算され、ランダムビット  $WT\_RBITS1115$  のテーブルが送信者と受信者との間で秘匿される場合、通信システム 100 は、安全な通信システムとして使用され得る。なぜなら、 $W(I)$  のシーケンスを知らずに、受信したストリームを復号化することは、不可能ではないが、困難であるからである。この特徴はまた、 $W(I)$  を計算または生成する方法である限り、 $W(I)$  を計算または生成する他の方法を使用する通信システムに適用し得る、また、好ましくは、発生器  $W(I)$  を計算する方法において使用されるシードは、秘匿される。

10

【0129】

(別のデコーダ)

この開示を読むと、受信器は、上記の実施例よりも効率的に動作し得ることが当業者に明らかである。例えば、受信器は、パケットをバッファし、1 群のパケットが到着したときだけ回復ルールを適用すればより効率的であり得る。この改変は、後の不必要な動作を行う際にかかる計算時間を低減し、コンテキストスイッチングによるオーバーヘッドを低減する。特に、デコーダは、少なくとも  $K$  個の出力記号 (同じサイズの入力記号および出力記号を仮定する) がパケット単位で到着する前に、 $K$  個の入力記号の元のファイルを回復することは望めないで、復号化プロセスを開始する前に少なくとも  $K$  個の出力記号が到着するまで待機することが有益であり得る。

20

【0130】

図 13 は、上記の概念を含み、図 6 のデコーダによって使用されるプロセスの改変例である復号化の異なる方法を示す。2 つの方法間の主な違いは、図 13 の方法が 1315 に示すように、バッチ単位で出力記号の群を受信することである。第 1 のバッチのサイズは、到着する出力記号の総数が  $K + A$  であるように設定され、ここで、 $A$  は、入力記号数  $K$  の小さな部分である (1310)。出力記号の群の第 1 のバッチが受信された後で、出力記号の群は、出力記号の群を処理するために受信オーガナイザ 520 を使用するステップ (1340) と、復号可能なセットを処理し、たかだかそれらの群サイズであり低減された重みを有する出力記号の群から入力記号を回復するために回復プロセッサ 525 を使用するステップ (1350) とを組み合わせ、前記のように処理される。 $K$  個の入力記号のすべてを回復することを、出力記号の群の第 1 のバッチを使用して達成できない場合、各バッチがさらなる  $G$  個の出力記号を含む、出力記号のさらなるバッチが、すべての入力ファイルが回復されるまで受信および処理される。

30

【0131】

デコーダの補助データ構造に必要な記憶装置をできるだけ低減することは、利点である。すでに説明したように、出力記号の各群に対する関連付けのリストのための記憶装置は必要でない、なぜなら、必要に応じて、連想装置 220 を使用して、それらのリストを速やかに計算し得るからである。まだ回復されない入力記号の各々に対して、関連付けとして入力記号を有する出力記号の群の  $GOSDS505$  における行番号を格納するために別の記憶装置が必要である。すなわち、図 5 のテーブル  $ISDS510$  における  $RL$  列に示されるリストのためのスペースが必要である。図 8 のステップ 855 においてすでに説明したように、この格納を使用すると、所定の入力記号が再構築された場合、どの出力記号の群が低減可能かを速やかに識別し得る。それが効率良くなされなければ、これらのリストに必要な記憶装置は、すべての入力記号を回復するために使用される出力記号のすべての群の関連付けの総数に比例し得る。

40

【0132】

(予めソートするデコーダ)

次に、図 14 および 15 を参照して、デコーダのより好ましい実施形態を説明する。図 14 は、デコーダを構成する部分を示す。それらは、図 5 に示すものと同じであるが、ただ

50

し、テーブルWEIGHT SORT 1405および図8bにおいて説明されるように形成された実行スケジュールを格納するために使用されるEXECUTION LIST 1420が追加される点で異なる。出力記号の群のGOSDS 505における行番号のバッチを格納するために、テーブルWEIGHT SORTが使用され、それらは受信される際に、重みの昇順で格納される。WT\_\_VAL列を使用して、重みを格納し、ROW\_\_LIST列を使用して、GOSDS 505における出力記号の群の行番号を格納する。一般に、重みWT\_\_VAL(L)を有する出力記号のすべての群の行番号は、ROW\_\_LIST(L)に格納される。このテーブルを使用して、図15のステップ1520に示すように、重みの昇順で出力記号の群のバッチを処理する。出力記号の低い重み群は、入力記号を回復するためにそれほど集中して計算に使用せず、出力記号の群の重みが大きいほど、それらの関連付けのほとんどがただちに回復されるということがあり得る。したがって、それは、リンク記憶装置スペース(デコーダは、回復された入力リンクによって使用されるスペースを回復し得、処理中の出力記号の群はまだ未回復のわずかな関連付けを有する)を実質的に節約する。

#### 【0133】

出力記号の群を適切なサイズのバッチで重みの昇順に処理することは、メモリ要件および処理要件を下げる。

#### 【0134】

図15に示すように、出力記号の群においてK個よりわずかに多い出力記号(図ではK + A個の出力記号によって示す)は、いずれの処理の開始よりも前に到着することが許される(1515)。ここで、入力ファイルにおける同じサイズの入力ファイルおよび出力ファイル、ならびにK個の入力記号を仮定する。初めに、デコーダは、群におけるK + A個の出力記号の受信を単に待機する。なぜなら、デコーダは、K + A個の出力記号より少ない出力記号から入力ファイルをどうしても回復し得ないと予想でき、おそらくK個より少ない出力記号から任意の入力ファイルを回復し得ない。実用上は、5・KがAに対して良好な値であることが分かった。

#### 【0135】

出力記号の受信された群のGOSDS 505における行番号は、図14のテーブルWEIGHT SORT 1405において重みの昇順で格納される(図15のステップ1515)。Tが、そのときの1とTとの間のLの値に対する出力記号の重みの可能な群の数である場合、リストROW\_\_LIST(L)は、重みWT\_\_VAL(L)の出力記号の受信されたすべての群を含み、ここで $1 = WT\_VAL(1) < WT\_VAL(2) < WT\_VAL(3) < \dots < WT\_VAL(T)$ であり、かつWT\_\_VAL(T)が出力記号の任意の群のうちの最大の重みである。次に、ステップ1520に示すように、図15に示すデコーダの動作の残りは、図13に示すデコーダとまったく同じであるが、ただし、出力記号の群が重みの昇順で処理されることを除く。

#### 【0136】

通常、K + A個の出力記号は、すべての入力記号を回復するのに十分であり得る。しかし、全部でK + A個の出力記号を含む出力記号のいくつかのセット群は、すべての入力記号を回復するのに十分でないことがある。そのような場合、G個のさらなる出力記号のバッチは、群単位で受信され、そしてすべての入力記号が回復されるまで処理される。Gに対する良好な設定は、Kである。

#### 【0137】

図16~19は、図15において記載するプロセスの実行例の1スナップショットを示す。この例において、群サイズは、すべての群に対して2であり、出力記号の4個の群(16030、16035、16040、16045)は、図16において矢印付きの線で示す関連付け(16000、16005、16010、16015、16020、16025)とともに受信されている。まず、キー23および値(A, D)、キー159および値(RS\_\_1(A, E, F), RS\_\_2(A, E, F))、キー33および値(B, C)、ならびにキー835および値(RS\_\_1(C, F, G, H), RS\_\_2(C, F, G,

10

20

30

40

50

H))を有する出力記号の群(第1および第3の出力記号の群は、重み2を有し、その値関数が識別関数であり、第2および第4の群は、それぞれ重み3および重み4を有し、その値関数がリード・ソロモン値関数である)が、図17に示すようにG O S D S 5 0 5に受信および格納される。G O S D S 5 0 5における行番号は、図18に示すように、出力記号の重みに対応する行においてROW\_\_LIST中に格納される。重み2の出力記号の群は、G O S D S 5 0 5の行0および行2にある。このように、ROW\_\_LIST(0)は、重みWT\_\_VAL(0)=2を有する出力記号の群に対応し、図18に示すように、行番号0および2を含む。同様に、WT\_\_VAL(1)=3であり、かつROW\_\_LIST(1)は1を含み、WT\_\_VAL(2)=4であり、かつROW\_\_LIST(2)は3を含む。

10

#### 【0138】

プロセスのこの時点で、重み昇順にした出力記号の最初の2つの群が処理されており、G O S D S 5 0 5の行1における出力記号の第3の群は、受信オーガナイザ520によって処理されており、出力記号のこの第3の群が回復プロセッサ525によってちょうど処理されるところである。行0および2における出力記号の群が、スケジュールにすでに付加され、それぞれ位置0、3、1および2で最終的に入力記号を回復する。G O S D S 5 0 5の行3における出力記号の群は、まだ回復されていない位置5、6および7の3つの関連付けを有し、したがって、I S D S 5 1 0における位置5、6および7からG O S D S 5 0 5における行3に戻るリンクが存在する。G O S D S 5 0 5の行1における出力記号の群は、位置0、4、および5において3つの関連付けを有する。位置0における関連付けは、回復されたものとしてすでにマークされており、したがって、そこから行1に戻るリンクは存在しない(それにより、出力記号のこの群の重みが3から2に低減され、このことは、一旦回復プロセッサ525が実行されると位置4、5、6および7における残りの入力記号の回復をトリガする)。位置4および5における関連付けは回復されず、したがって受信オーガナイザ520は、I S D S 5 1 0における位置4および位置5からG O S D S 5 0 5における行1へのリンクを付与する。これはすべて、図17に示される。このように、プロセスのこの時点で、入力記号から、それらを関連付けとして有する出力記号の群へ戻る全部で5つのリンクだけが使用される。これは、あらゆる入力記号から、それを関連付けとして有するあらゆる出力記号の群へのリンクを使用する単純な実施例に遜色がない。この例において、このようなリンクが11可能である。

20

30

#### 【0139】

一般に、リンク用の格納スペースの節約が、図13に記載したプロセスよりも図14および15において記載したプロセスを使用する場合、劇的に低減する。例えば、入力記号数が50,000の場合、スペースの節約は、一般にリンクスペースにおいて10~15倍である。この低減の理由は、重みのより小さい出力記号の群は、プロセスの終了時より開始時に入力記号を回復する可能性がより大きく、重みのより大きな出力記号の群は、プロセスの開始時より終了時に出力記号の群を回復する可能性がずっとより大きい。したがって、重みの昇順で出力記号の群を処理することは意味のあることである。図13よりも図14および15に記載するプロセスのさらなる利点は、復号化が一般に30%~40%速くなることである。これは、重みのより小さい出力記号の群が、重みのより大きい出力記号の群よりも、入力記号を回復するために使用される可能性が大きく(なぜなら、重みのより小さい出力記号の群が先に考慮されるからである)、かつ特定の入力記号を回復するコストは、それを回復するために使用される出力記号の群の重みに直接依存するからである。

40

#### 【0140】

図19は、この例の完全な実行スケジュールを示し、実行時に、群サイズ2を有する出力記号の4個の群に基づいて8個の入力記号すべてを回復し得る。

#### 【0141】

(重み分布の選択)

重要な最適化は、入力ファイルをできるだけ少ない出力記号の群を用いて完全に再構築し

50

得るように符号化プロセスを設計することである。この最適化は、送信時間および帯域幅がコスト高または制限される場合、または入力ファイルが、さらなる出力記号の群を待つことなく速やかに復号化されなければならない場合、役に立つ。一般に、入力ファイルの再構築に必要な十分な出力記号数は、元の入力ファイルにおける入力記号数よりもわずかに大きい（同じサイズの入力記号および出力記号を仮定する）。入力ファイルよりも少ないビットを受信した場合、任意の入力ファイルは回復され得ないことを示し得る。したがって、完全な符号化スキームは、入力ファイルと同じビット数を符号化する任意の出力記号の群のセットから入力ファイルを回復することが可能であり得、符号化効率の1つの尺度は、予期される条件下に必要な予備のビットがどれくらい少ないかである。

#### 【0142】

図5に示すデコーダにおいて、最大効率は、デコーダがちょうどK個の出力記号を受信した後で回復プロセッサ525が最後の未知の入力記号を回復する場合に得られる。K個より多い出力記号が、すべての入力記号が回復され得る時間までにデコーダによって受信されたならば、入力ファイルの回復に必要でない、または使用されない出力記号の群が受信されたであろう。最大効率が良好である間、それを目標とすることは、再構築が完全になる前にDSS515が空になり得る危険によって調節されるべきである。言い換えると、最大効率において、復号可能なセットのサイズは、ちょうど再構築が終了する際にゼロとなるが、符号化/復号化は、復号可能なセットのサイズが再構築の終了前にゼロとなるわずかの確率しかないように、 $K + A$ 個の出力記号を使用して調整されるべきであり、そのため群に含まれるG個の出力記号のさらなるセットは必要でない。

#### 【0143】

この点を図20に例示する。この図は、復号可能なセットのサイズ対再構築された入力記号数のプロットを示す。ここで、デコーダは、以下に記載の理想の分布に対して、出力記号の $K/b$ 個の群を用いて動作し、各群はサイズbを有する。この例において、 $A = 0$ 、すなわち、K個の入力記号のすべてを復号するために、群に受信された出力記号数は、可能な数の最小数である（入力記号および出力記号は同じサイズと仮定する）。理解すべきことは、そのプロットが、重みおよび関連付けを決定するための任意の所定関数に対して異なり、また出力記号のどの特定の群を受信したかに依存して異なり得ることである。そのプロットにおいて、復号可能なセットサイズの予想サイズは、最初は1であり、回復プロセスを通して1を維持する。このように、予想される挙動においては、次のb個の入力記号を回復するために使用され得る、復号可能なセットにおける出力記号の1群が常に存在する。図20はまた、理想の分布の実際の挙動の例を示す。この実際の実行において、復号可能なセットは、回復が完了する前は、すなわち、100, 000個の入力記号のうちの2つのみが回復された後は、空であることに留意されたい。理想の分布の実際の挙動は、一般的である、すなわち、理想の分布に対して、ランダムなゆらぎはほとんど常に、すべての入力記号が回復される前に復号可能なセットを空にし、このため、以下に記載のように、よりロバストな分布が必要である。

#### 【0144】

効率は、復号可能なセットにおける出力記号の群が、関連付けの値を再構築するために使用される場合に、その群サイズより厳密に小さい未知の入力記号を関連付けとして有する回数を限定することによって、向上される。これは、 $W(I)$ を生成するための関数を適切に選択することによって達成され得る。

#### 【0145】

このように、出力記号の十分な群を受信することによって、入力ファイルを任意の所望の確実度で完全に回復することが可能である一方で、Aの何らかの小さい値に対して $K + A$ 個程度に少ない出力記号（入力記号および出力記号は同じサイズと仮定する）を用いて完全な入力ファイルを含むK個の入力記号を回復する確率が高くなるように群連鎖反応符号化通信システムを設計することが好ましい。Aの最小値がゼロであり、各出力記号がK個の入力記号のすべてに依存する標準的なリード-ソロモン符号化を使用する場合などのいくつかの符号化スキームにおいて達成され得るが、これにより、符号化および復号化時間

が遅くなる。Aとして何らかの小さい非ゼロ値を許容することによって、改善された通信システムが得られ得る。

#### 【0146】

小さい値のAは、群サイズ、出力記号の群に対する重み分布（すなわち、すべてのIについての $W(I)$ の分布）、および出力記号の群についての関連付けの分布（すなわち、すべてのIについての $AL(I)$ のメンバーシップ）を決定するための適切な分布を使用することによって、群連鎖反応符号化において達成され得る。強調すべきことは、復号化プロセスが群サイズ、重み分布および関連付けの選択についての分布にかかわらず適用され得るが、好ましい実施形態は、群サイズ、重み分布および特に最適に近い性能として選択された関連付けの選択についての分布を使用し得る。実際、選択した分布の小さな変化によって性能はほんのわずかしかならば変化し得ないので、多くの分布が良好に機能し得る。

10

#### 【0147】

1つの好ましい実施形態による分布を決定するための方法論をここで説明する。この実施形態において、出力記号のすべての群の群サイズは、固定された正の整数bと同じ値bである。使用される実際の重み分布は、理想の数学的分布に基づく。理想の重み分布において、重み $W(I)$ は、「理想」の確率分布にしたがって選択される。最小の重みはbであり、最大の重みはKである。ここで、Kは、入力記号数である。理想の分布において、iの値に等しい重みが以下の確率pを用いて選択される：

$i = b$  の場合、 $p = b / K$  ; および

$i = b + 1, \dots, K$  の場合、 $p = b / (i(i - 1))$

20

一旦重み $W(I)$ が選択されると、 $W(I)$ 個の関連付けられた入力記号のリスト $AL(I)$ がランダムに（必要ならば、擬似ランダムに）独立かつ均一に選択され、選択された関連付けの全てが確実に異なるようにする。このように、第1の関連付けは、K個の入力記号からランダムに選択される（各入力記号は、選択される確率が $1/K$ である）。次に、第2の関連付け（ $W > 1$ の場合）は、残りの $K - 1$ 個の記号からランダムに選択される。上記の重み確率分布は、システムが予想通り正確に挙動するならば、正確に出力記号の $K/b$ 個の群がすべての入力記号を復号および回復するのに十分であり得るという性質を有する。理想の分布に対するこの予想される挙動は、図20において実線で示される。しかし、重みおよび関連付けの選択のランダム性質のために、かつ出力記号の群の任意のセットが復号化プロセスにおいて使用されるために、プロセスは、必ずしもそのように挙動し得ない。理想の分布に対する実際の挙動の例を図20において点線で示す。このように、理想の重み分布は、実用上はいくらか改変されなければならない。

30

#### 【0148】

一般に、所定の設定に対する最良のパラメータは、コンピュータシミュレーションによって見つけられ得る。しかし、理想の重み分布に関する簡単な変形例は、効果的なオプションである。この簡単な変形例において、理想の重み分布は、重みbを有する出力記号の群および出力記号の大きな重みを有する群の確率を上げることによってわずかに改変されて、K個すべての入力記号が回復される前に復号可能なセットが空になる確率を低減する。出力記号の重みbを有する群を余分に供給すると、回復プロセスが入力記号の回復の終了近くになるまで、そのプロセスが出力記号の重みがたかだかbである群を使い果たす（すなわち、復号可能なセットを空にする）確率を低減する。出力記号の大きな重みの群を余分に供給すると、回復プロセスの終了近くで、各々まだ回復されない入力記号に対して、関連付けとしてその入力記号およびまだ回復されない関連付けとしてたかだか $b - 1$ 個の他の入力記号を有し得る出力記号の少なくとも1つの群が存在する確率を増加する。

40

#### 【0149】

より詳細には、改変された重み分布は以下の通りである：

$i = b$  の場合、 $p = n \cdot b \cdot R1 / K$  ;

$i = b + 1, \dots, (K / R2 - 1)$  の場合、 $p = n \cdot b / (i(i - 1)(1 - i R2 / K))$  ; および

$i = K / R2, \dots, K$  の場合、 $p = n \cdot HW(i)$

50

ここで、 $K$ は入力記号数であり、 $R_1$ および $R_2$ は調節可能なパラメータ、ならびに $n$ は $p$ 値のすべての合計が1となるように使用される正規化因子である。

#### 【0150】

$HW(i)$ ならびに $R_1$ および $R_2$ としてサンプル値の計算を、添付書類Aにおいて詳細に示す。そこで、C++の記号`nStartRippleSize`、`nRippleTargetSize`および`nSymbols`は、それぞれ上記式の $R_1$ 、 $R_2$ および $K$ に対応する。

#### 【0151】

この改変された分布は、理想の数学的重み分布に同様であり、重み $b$ およびさらに大きな重みを有する出力記号のさらに多くの群、ならびにそれにしたがって再設計された分布を有する。改変された分布において示されるように、 $R_1$ は、重み $b$ の記号の複数の増加した確率を決定すると同様に、出力記号の重み $b$ の群の初期の割合を決定し、 $R_2$ は、「より大きな」重みと「それほど大きくない」重みとの間の境界を決定する。

#### 【0152】

$R_1$ および $R_2$ に対する良好な選択は、一般に $K/b$ に依存し、経験的に決定され得る。例えば、 $R_1$ を $1.4 \times (K/b \text{の平方根})$ に等しくし、 $R_2$ を $2 + 2.1 \times (K/b \text{の四乗根})$ に等しくすると、実用上うまくいく。したがって、 $K = 4000$ および $b = 5$ の場合、 $R_1 = 3.9$ および $R_2 = 1.3$ に設定するとうまくいく。 $K$ が $64000$ および $b = 2$ の場合、 $R_1 = 2.50$ および $R_2 = 3.0$ に設定するとうまくいく。 $R_1$ および $R_2$ のより詳細な計算を、添付書類Aに示す。図21は、この分布の予想される挙動が、回復プロセス全体にわたって復号可能なセットを適度に大きくしたままにするので、実際の実行下では、予想される挙動からのランダムなゆらぎは、すべての入力記号が回復される前に復号可能なセットを空にする見込みはない。重み分布の適切な設定によって、いかなる損失条件下においても、最小数の出力記号を受信すれば高い確率で入力記号のすべてを回復することを確実なものとする。

#### 【0153】

上記の再構築プロセスは、トルネード符号に対して使用されるものと同様であるが、符号を形成するために使用される異なるプロセスによって、極端に異なる効果を生じる。特に、上記のように、群連鎖反応符号化に必要なメモリは、トルネード符号に必要なメモリよりも著しく小さく、種々の状況における群連鎖反応符号化の使用の容易さは、いくらかのスピードを犠牲にする可能性はあるが、トルネード符号のそれをはるかに超える。そのプロセスの基礎をなす数学的詳細は、以下により詳細に説明される。

#### 【0154】

(いくつかの群連鎖反応符号の性質)

生成され、そしてチャネルを介して送信される出力記号の群の数は、他の符号化スキームと同様に、群連鎖反応符号化を用いても制限されない。なぜなら、キーは、入力記号に対して1対1対応を有する必要がなく、 $I$ の異なる値の数は、入力記号数の何らかの小さな一定の割合に制限されないからである。したがって、復号可能セットが、入力ファイルが再構築される前に空になる場合でさえ、復号プロセスが失敗しない可能性がある。なぜなら、デコーダが、必要とされるより多くの出力記号の群を集めて、重みはその群サイズである出力記号の群を少なくともさらに1つ得ることができる。その群サイズの重みの出力記号からなるその群が受信されると、それは、復号可能なセットに置かれ、連鎖反応効果によって、前に受信された出力記号の群をたかだかそれらの群サイズの重みに低減させるので、それらは入力記号の再構築のために使用され得る。

#### 【0155】

上記の例のほとんどにおいて、入力および出力記号の群は、同じ数のビットとして符号化し、出力記号の各群は、1つのパケットに配置される(1パケットは、送信の1単位であり、完全に受信されるか、または完全に損失される)。いくつかの実施形態において、通信システムは、各パケットが出力記号の数個の群を含むように改変される。次に、出力記号値のサイズは、(変化し得る)群サイズを含む多くのファクタに基づいて、最初にファ

イルを入力記号に分割した際の入力記号のサイズによって決定されるサイズに設定される。符号化プロセスは、実質的に変更されないままであり得る。ただし、出力記号の群が、各パケットが受信される際に、ひとまとまりで到着し得ることを除く。

#### 【0156】

入力記号および出力記号サイズの設定は、通常、ファイルのサイズおよび出力記号の群が送信される通信システムによって決定される。例えば、通信システムは、データのビットを規定サイズのパケットにグループ化するか、または他の方法でビットをグループ化する場合、記号サイズの設定は、パケットまたはグループ化サイズから開始される。そこから、設計者は、出力記号の何個の群を1つのパケットまたは群で送信するか、かつ（変化し得る）その群サイズとともに出力記号サイズを決定し得る。簡単のために、設計者は、おそらく入力記号サイズを出力記号サイズと等しくなるように設定し得るが、入力データが異なる入力記号サイズをより便利にする場合は、それが使用され得る。

10

#### 【0157】

入力記号サイズを決定する際の別のファクタは、入力記号数  $K$  が受信オーバーヘッドを最小に保つために十分大きくなるように入力記号サイズを選択することである。例えば、 $K = 10,000$  にすると、平均の受信オーバーヘッドが、5% ~ 10% の適度なゆらぎを有するようになり、他方  $K = 80,000$  にすると、平均の受信オーバーヘッドは1% ~ 2% で、そのゆらぎが非常に小さくなる。例として、 $K = 80,000$  で1,000,000回の試行を含む1つのテストにおいて、受信オーバーヘッドは、4%を決して超えなかった。

20

#### 【0158】

上記符号化プロセスは、元のファイルに基づき、出力記号の群を含むパケットの1ストリームを生成する。出力記号の群は、ファイルの符号化形態またはより簡潔には符号化ファイルを保持する。ストリーム中の出力記号の各群は、出力記号の他のすべての群と独立して生成され、生成され得る出力記号の群の数には下限および上限がない。キーは、出力記号の各群に関連付けられる。そのキーおよび入力ファイルの内いくつかの内容は、出力記号の群の値を決定する。出力記号の連続生成された群は、連続するキーを有する必要はなく、いくつかのアプリケーションにおいては、キーのシーケンスをランダムに生成するか、またはシーケンスを擬似ランダムに生成することが好ましくあり得る。

30

#### 【0159】

群連鎖反応符号化は、元のファイルが  $K$  個の等サイズの入力記号に分割され得、かつ各出力記号値が入力記号値と同じ長さを有する場合に、そのファイルが平均で  $K + A$  個の出力記号から回復され得る（ここで  $A$  は  $K$  に比較して小さい）という性質を有する。例えば、 $A$  は、 $K = 20,000$  の場合500であり得る。出力記号の特定の群がランダムまたは擬似ランダムな順序で生成され、かつ送信中の出力記号の特定の群の損失が任意に仮定されるので、入力ファイルを回復するために必要な出力記号の群の実数の数にはいくらかの小さなばらつきが存在する。 $K + A$  個の出力記号を含む特定の集合パケットが、入力ファイル全体を復号するのに十分でないようないくつかの場合において、受信器が1つ以上のパケットのソースからより多くのパケットを集め得る場合は、入力ファイルは、依然として回復可能である。

40

#### 【0160】

出力記号の群の数は、 $I$  の解像度によってのみ制限されるので、 $K + A$  個より十分に多くの出力記号の群が生成され得るべきである。例えば、 $I$  が32ビット数である場合、四十億個の異なる出力記号の群が生成され得る。他方、そのファイルは、 $K = 50,000$  個の入力記号を含み得る。実用上は、これら四十億個の出力記号の群のうちのほんのわずかな個数の群だけが生成および送信され得、そして入力ファイルが、出力記号の可能な群の非常に小さな割合を用いて回復され得ることがほぼ確実であり、そして入力ファイルが、 $K$  個よりわずかに多くの出力記号を用いて回復され得る可能性は非常に高い（入力記号サイズは、出力記号サイズと同じであると仮定する）。

#### 【0161】

50

出力記号の各群を生成するために必要な算術演算の平均数は、 $\log K$  に比例し、そのため、入力ファイルを復号および回復するために必要な算術演算の総数は、 $K \log K$  に比例する。上に示すように、入力ファイルを格納するために必要なメモリよりもほんのわずかに多くのメモリを使用する効率的な復号化プロセスが存在する（一般に、約 15% 多い）。上記の数は、従来から公知の符号化技術と比較して演算および記憶装置において著しい減少を示す。

#### 【0162】

例えば、標準的なリード・ソロモン符号は、通信アプリケーションのための標準的な符号である。標準的なリード・ソロモン符号を用いると、群連鎖反応符号化を用いるように、入力ファイルは  $K$  個の入力記号に分割されるが、リード・ソロモン符号におけるその  $K$  個の入力記号は、 $N$  個の出力記号に符号化される。ここで、 $N$  は一般に、符号化プロセスが開始する前に固定される。これは、出力記号の群の不確定数を考慮する本発明と対照的である。

10

#### 【0163】

出力記号の群の不確定数を有する 1 つの利点は、受信者が予想より多くの出力記号の群を逸する場合、不十分なチャネルによるか、または出力記号のいくつかの群がすでに通過した後で開始する受信者により、受信者は、単に少し長く受信し、出力記号のより多くの群を取り上げ得ることである。別の利点は、受信者が複数のエンコードから生成される出力記号の群を集め得るので、各エンコードは、入力記号を復号するために必要な出力記号の群のほんの小さな割合だけを提供する必要があり、1 つのエンコードからの出力記号の群の数は、どれだけのエンコードが出力記号の群を受信者に供給しているかに依存し得ることである。

20

#### 【0164】

標準的なリード・ソロモン符号はまた、符号化および復号化の両方について、連鎖反応符号より実質的に多くの時間を必要とする。例えば、標準的なリード・ソロモンを用いて各出力記号を生成するために必要な算術演算の数は、 $K$  に比例する。標準的なリード・ソロモン符号を復号化するために必要な算術演算の数は、どの出力記号が受信者に到着したかに依存するが、一般に、そのような演算の数は、 $(N - K) \cdot K$  に比例する。このように、実用上は、 $K$  および  $N$  の許容値は、非常に小さく、数十のオーダー、おそらくわずか数百までのオーダーである。例えば、クロス・インターリーブ (Cross-Interleaved) リード・ソロモン符号は、コンパクトディスク (CD) および CD-ROM 上で使用される。CD の場合、1 つの標準的な符号は  $K = 24$  および  $N = 28$  を使用し、別の標準的な符号は  $K = 28$  および  $N = 32$  を使用する。CD-ROM の場合、1 つの標準的な符号は  $K = 24$  および  $N = 26$  を使用し、別の標準的な符号は  $K = 43$  および  $N = 45$  を使用する。MP EG ファイル (MP EG は、ビデオおよびオーディオストリームのためのファイルフォーマットである) の衛星送信のために使用される標準的なリード・ソロモン符号は、 $K = 188$  および  $N = 204$  を使用し、一般に、このような大きな値は、専用のハードウェアを必要とする。

30

#### 【0165】

$c K \log K$  時間での符号化および  $c' K (\log K)^2$  時間での復号化を可能にする標準的なリード・ソロモン符号のより高速な実施例が存在することが公知であるが、 $c$  および  $c'$  が過度に大きな定数であるので、 $K$  の非常に大きな値以外全てに対しては、これらの実施例は標準的なリード・ソロモン符号の他の実施例よりも遅くなる。すなわち、数千または数万の  $K$  の値に対して有効なクロスオーバーポイントが存在する。したがって、クロスオーバーポイントより下の  $K$  の値の場合、標準的なリード・ソロモン符号の他の実施例は、より高速である。そのより高速な実施例は、クロスオーバーポイントより高い  $K$  の値の他の標準的なリード・ソロモン符号よりも高速であるが、そのより高速な実施例は、 $K$  のそれらの値で群連鎖反応符号よりも数桁遅い。

40

#### 【0166】

その速度制限のために、標準的なリード・ソロモン符号に対しては、一般に、小さい値の

50

KおよびNだけが実現可能である。したがって、大きなファイルに関してそれらを使用することにより、そのファイルを多くのサブファイルに分割し、各サブファイルを別々に符号化することを必要とする。そのような分割は、符号の有効性を低減し、送信中のパケット損失を防止する。

#### 【0167】

標準的なリード・ソロモン符号の1つの特徴は、K個の異なる出力記号のいずれもが、入力ファイルを復号化するために受信者によって使用され得ることである。おそらく、少なくともK個の出力記号が、任意の入力ファイルを復号化するために必要とされ、したがって、標準的なリード・ソロモン符号はこの点で最適であるということが証明できる。なぜなら、Kはまた、入力ファイルを復号化するために必要な出力記号の最大数であるからである。対照的に、群連鎖反応符号化は一般に、 $K + A$ 個の出力記号（Aは、適切に選択されたKに比較して小さい）または合計でK個よりわずかに多い出力記号を必要とする。上記のネットワークアプリケーションにおいては、おそらくA個のさらなる記号が必要であるというこの欠点は、速度の利点および途切れなくより大きなファイル进行处理する能力によって極めて影を薄くしている。

#### 【0168】

（基本通信システムの変形例）

1つのチャネルに対する本発明による通信システムの実施形態は、上記に詳細に説明された。これらの実施形態の要素は、1つより多いチャネルの有利な使用に拡張し得る。

#### 【0169】

図22～23は、図1に示すような通信システムを組み込んだ2つのコンピュータの間のシステムを示す。第1の例（図22）は、入力ファイル2210を受信者コンピュータ2220にネットワーク2230を介して送信する送信者コンピュータ2200を有する。第2の例（図23）は、入力ファイル2310を受信者コンピュータ2320（図には1つだけを示す）へ無線チャネル2330を介して放送する送信者コンピュータ2300を有する。ネットワーク2330の代わりに、インターネットのワイヤなどの任意の他の物理的通信媒体を使用し得る。無線チャネル2330は、無線ラジオチャネル、ページャーリンク、衛星リンク、または赤外線リンクなどであり得る。図23に示す構成はまた、1送信者が入力ファイルを多くの受信者に送信する場合、受信者が入力ファイルを多くの送信者から得る場合、または多くの受信者が入力ファイルを多くの送信者から得る場合に、有線または無線媒体とともに使用され得る。

#### 【0170】

上記説明を読めば明らかなように、上記の群連鎖反応符号化スキームを使用して、コンピュータネットワーク、インターネット、モバイル無線ネットワークまたは衛星ネットワークなどの損失のある送信媒体を介して、ファイルを所望の性質を有するパケット化されたデータのストリームとして送信し得る。そのような所望の性質の1つは、復号化エージェントは、一旦そのストリームから十分に多くのパケットの任意のセットを受信すると、元のファイルを極めて速やかに再構築し得ることである。群連鎖反応符号化はまた、マルチキャストまたは放送設定において1送信エージェントが、そのファイルを複数の受信エージェントに送信する場合などの、多くのエージェントが関与する状況において有用である。

#### 【0171】

群連鎖反応符号化スキームはまた、複数の送信エージェントが同じファイルを複数の受信エージェントに送信する状況においても有用である。これにより、ネットワークインフラストラクチャの局所的な失敗に対して向上したロバスト性を可能にし、受信エージェントが受信するパケットの送信エージェントを1送信エージェントから別の送信エージェントに途切れなく変更でき、かつ受信エージェントが同時に1つより多い送信エージェントから受信することによってそれらのダウンロードを高速化できる。

#### 【0172】

1つの局面において、上記の群連鎖反応符号化プロセスは、ホログラフィーイメージのデ

10

20

30

40

50

デジタル等価物を実行する。ここで、送信の各部分は、送信されるファイルのイメージを含む。ファイルがメガバイトのファイルである場合、ユーザは、ただ送信されるストリームを利用していずれの任意のメガバイトに値するデータ（それに加えていくらかの余分のオーバーヘッド）を得、そのメガバイトから元のメガバイトファイルを復号し得る。

#### 【0173】

群連鎖反応符号化は、送信されたストリームからデータをランダムに選択して動作するので、ダウンロードは、スケジュールされる必要はなく、一貫している必要もない。群連鎖反応符号化を用いたビデオ・オン・デマンドの利点を考察する。特定のビデオが、受信器と送信器との間で調整することなく特定のチャンネル上を連続するストリームとして放送され得る。受信器は、興味のあるビデオ用の放送チャンネルに単に同調し、送信がいつ開始したかとか、または放送の損失部分のコピーをどのように得るかを理解する必要なしに、元のビデオを再構築するために十分なデータをキャプチャする。

10

#### 【0174】

これらの概念を、図24～25に例示する。図24は、1つの受信器2402が3つの送信器2404（それぞれ、「A」、「B」および「C」で示す）から3つのチャンネル2406を介してデータを受信する構成を例示する。この構成を使用して、受信器が利用可能な帯域幅を3倍にし得るか、またはいずれか1つの送信器からファイル全体を得るための十分な長さで利用可能でない送信器を扱い得る。図示のように、各送信器2404は、値のストリーム $S(I)$ を送信する。各 $S(I)$ 値は、キー $I$ および出力記号値の1群 $B(I) = B_{-1}(I), \dots, B_{-b}(I)$ を表し、その使用については上記している。例えば、値 $S(n_A + i_A n'_A)$ は、値 $n_A + i_A n'_A$ を有するキーであり、それに値 $B(n_A + i_A n'_A)$ を有する出力記号の群が続く。ここで、 $n_A$ および $n'_A$ は、送信器2404（A）に関連付けられたランダムに選択された数であり、 $i_A$ は、送信器2404（A）から送信されるパケットのシーケンス番号である。1つの送信器からのキーのシーケンスは、好ましくは他の送信器からのキーのシーケンスと異なっているので、送信器は誤りを重複させない。これは、各送信器で使用するキーシーケンスがその送信器に特定の情報の関数である（例えば、送信器2404（A）のキーシーケンスは $n_A$ および $n'_A$ に依存する）という事実によって図24に例示する。

20

#### 【0175】

送信器2404は、作業を重複させないように同期または調整される必要がないことに留意されたい。実際、調整することなく各送信器は、大きく異なるキーのシーケンスを送信する可能性がある（すなわち、 $n_A + i_A n'_A, n_B + i_B n'_B, n_C + i_C n'_C$ 、ここで、 $i_A, i_B, i_C$ は、それぞれ衛星SAT A、SAT B、SAT Cによって送信されたパケットのシーケンス番号であり、 $n_A, n'_A, n_B, n'_B, n_C, n'_C$ は、それぞれSAT A、SAT B、SAT Cに関連付けられたランダムに選択された数である）。

30

#### 【0176】

ランダムに選択した $K/b + A$ 個の出力記号の群、おそらくそれに伴うこと $G$ 個の余分の出力記号の群の数束を使用して、入力ファイルが回復され得るので、調整されない送信は、重複的である代わりに付加的である。

40

#### 【0177】

この「情報付加性」を図25に再度例示する。そこで、1つの入力ファイル2502のコピーが複数の送信器2504に提供される（そのうちの2つを図に示す）。送信器2504は、入力ファイル2502の内容から生成された出力記号の群をチャンネル2506を介して受信器2508に独立に送信する。各送信器が記号生成のために異なるセットのキーを使用する場合、そのストリームは、重複的ではなく独立かつ付加的（すなわち、それらは入力記号を回復するために使用される情報プールに付加する）である可能性がある。図示の2つの各送信器は、受信器のデコーダが入力ファイル全体を回復する事ができる前に（ $K/b + A$ ）/2個の出力記号の群を送信することだけが必要であり得る。

#### 【0178】

50

2つの受信器および2つの送信器を使用して、受信器部2510によって受信される総情報量は、1つのチャンネル2506を介して利用可能な情報の4倍の大きさであり得る。情報量は、例えば、送信器が両方の受信器に同じデータを放送する場合は、1つのチャンネルの情報の4倍より少なくてもよい。その場合、受信器部2510での情報量は、特にデータがチャンネルにおいて損失される場合、元のファイルを直接送信する事によって達成される速度の少なくとも2倍である。送信器が1個の信号しか放送しないが、受信器が異なる時間に見える所にある場合は、各送信器を受信する1つより多い受信器を有する利点がある。図25において、受信器部2510は、図1に示す受信器150、デコーダ155、キー発生器160および入力ファイルリアセンブラ165の機能と同様の機能を実行する。

10

#### 【0179】

いくつかの実施形態において、入力ファイル2502は、2つのエンコーダを有する1つの計算デバイスにおいて符号化されるので、計算デバイスは、1つの送信器に対して1つの出力を、他方の送信器に対して別の出力を提供し得る。これらの例の他の変形例は、この開示を読む際に明らかとなる。

#### 【0180】

本明細書中に記載の符号化装置および方法はまた、他の通信状況において使用され得、インターネットなどの通信ネットワークに制限されない。例えば、コンパクトディスク技術もイレーズおよび誤り訂正符号を使用して、傷の入ったディスクの問題に対応し、そこに情報を格納する際に群連鎖反応符号の使用により利益を受け得る。別の例として、衛星システムは、送信のためのパワー要件をトレードオフするためにイレーズ符号を使用し得、パワーを低減することによってより多くの誤りを意図的に考慮する。このアプリケーションにおいて、群連鎖反応符号化は有用であり得る。また、イレーズ符号は、情報記憶の信頼性のためのRAID（独立ディスクの冗長配列）システムを開発するために使用されてきた。したがって、本発明は、符号を使用して損失または誤りの可能性のあるデータの問題に対応する、上記の例などの他のアプリケーションにおける有用性を証明し得る。

20

#### 【0181】

いくつかの好ましい実施形態において、上記通信プロセスを実行するための命令のセット（またはソフトウェア）が、損失の可能性のある通信媒体を介して通信する2つ以上の多目的計算機に提供される。機械の数は、1送信者および1受信者から任意の数の送信機および/または受信機の範囲であり得る。機械を接続する通信媒体は、有線、光学、または無線などであり得る。上記通信システムは、本記載から明らかとなるべき多くの使用を有する。

30

#### 【0182】

上記は、例示的であって、限定的ではない。本発明の種々の変形例は、当業者にとって、本開示を読むことによって明らかとなり得る。したがって、本発明の範囲は、上記を参照して決定されるのではなく、その代りに添付の特許請求の範囲およびその等価物の完全な範囲を参照して決定されるべきである。

#### 【0183】

（結論として本発明は以下を提供する）

40

エンコーダは、データの入力ファイルおよびキーを使用して、出力記号の1群を生成する。キーIを有する出力記号の1群は、生成されるべきその出力記号の群の重み $W(I)$ を決定し、Iの関数にしたがってその群と関連付けられる $W(I)$ 個の入力記号を選択し、選択された $W(I)$ 個の入力記号の所定の値関数 $F(I)$ から出力記号値 $B(I)$ を生成することによって生成される。エンコーダは、繰り返し呼び出されて複数の出力記号の群または複数の出力記号を生成し得る。出力記号の群は、一般に互い独立であり、制限のない数（Iの解像度には影響される）が、必要に応じて生成され得る。デコーダは、生成された出力記号の一部またはすべてを受信する。入力ファイルの復号化に必要な出力記号の数は、ファイルを構成する入力記号の数に等しいか、またはわずかに大きい。ここで入力記号および出力記号は同じビット数のデータを表すと仮定する。図26a~eは、添付書

50

類 A の例示であり、重み分布を実施するためのプログラムのソースコードリストである。

【図面の簡単な説明】

【図 1】図 1 は、本発明の 1 つの実施形態による通信システムのブロック図である。

【図 2】図 2 は、図 1 のエンコーダをより詳細に示すブロック図である。

【図 3】図 3 は、出力記号の 1 群が関連付けられた入力記号の 1 セットからどのように生成され得るかの例示である。

【図 4】図 4 は、図 1 に示す通信システムにおいて使用され得るような基本デコーダのブロック図である。

【図 5】図 5 は、別のデコーダのブロック図である。

【図 6】図 6 は、出力記号の 1 セットの群から入力記号を回復するために、図 5 に示すデコーダなどのデコーダによって使用され得るプロセスのフローチャートである。

【図 7】図 7 は、出力記号の受信された群をオーガナイズするために、図 5 に示す受信オーガナイザなどの受信オーガナイザによって使用され得るプロセスのフローチャートである。

【図 8 a】図 8 a は、出力記号の受信された群を処理するために、図 5 に示す回復プロセッサなどの回復プロセッサによって使用され得るプロセスのフローチャートである。

【図 8 b】図 8 b は、図 8 a のプロセスの変形例の一部分のフローチャートであり、図 8 b は、保留される処理を含む回復プロセスにおいて実行されるステップを示す。

【図 8 c】図 8 c は、図 8 a のプロセスの変形例の一部分のフローチャートであり、図 8 c は、保留される処理を示す。

【図 9】図 9 は、図 2 の連想装置をより詳細に示すブロック図である。

【図 10】図 10 は、出力記号の群を用いて入力記号の関連付けを速やかに決定するために、図 9 に示す連想装置などの連想装置によって使用され得る 1 つのプロセスのフローチャートである。

【図 11】図 11 は、図 2 の重みセクタをより詳細に示すブロック図である。

【図 12】出力記号の所定の群の重みを決定するために、図 11 に示す重みセクタなどの重みセクタによって使用され得るプロセスのフローチャートである。

【図 13】図 13 は、特に効率的である必要のないデコーダのための復号化のプロセスのフローチャートである。

【図 14】図 14 は、より効率的なデコーダのブロック図である。

【図 15】図 15 は、図 12 ~ 13 を参照して説明される復号化よりも効率的に復号化するための、図 14 のデコーダを使用して実施され得るような復号化のためのプロセスのフローチャートである。

【図 16】図 16 は、図 15 の復号化プロセスのための、ドキュメントおよび受信された出力記号の群の例を例示する図である。

【図 17】図 17 は、図 15 に示す復号化プロセス中のデコーダにおけるテーブルの内容を例示する。

【図 18】図 18 は、図 15 に示す復号化プロセス中に使用され得るような重みソートテーブルの内容を例示する図である。

【図 19】図 19 は、図 15 に示す復号化プロセス中に形成され得る実行リストを例示する。

【図 20】図 20 は、理想の分布に対する、復号可能なセットサイズ対回復された入力記号数のプロットの形態の回復プロセスの進行を例示する。

【図 21】図 21 は、ロバストな重み分布に対する、復号可能なセットサイズ対回復された入力記号数のプロットの形態の回復プロセスの進行を例示する。

【図 22】図 22 は、上記図に例示するようなエンコーダおよびデコーダを使用して、1 送信者（送信器）と 1 受信器との間の 1 対 1 通信システムの例示である。

【図 23】図 23 は、上記図に例示するようなエンコーダおよびデコーダを使用して、1 送信者と複数の受信器（そのうちの 1 つだけを図示する）との間の放送通信システムの例示である。

10

20

30

40

50

【図 2 4】図 2 4 は、1 受信器が、複数の通常独立した送信者から出力記号の群を受信する場合の、本発明の 1 つの実施形態による通信システムの例示である。

【図 2 5】図 2 5 は、複数の独立であり得る受信器が、複数の通常独立した送信者から出力記号の群を受信して、1 受信器および / または 1 送信者だけが使用される場合より少ない時間で入力ファイルを受信する場合の、本発明の 1 つの実施形態による通信システムの例示である。

【図 2 6 a】図 2 6 a は、重み分布を実施するためのプログラムのソースコードリストである添付書類 A の例示である。

【図 2 6 b】図 2 6 b は、重み分布を実施するためのプログラムのソースコードリストである添付書類 A の例示である。

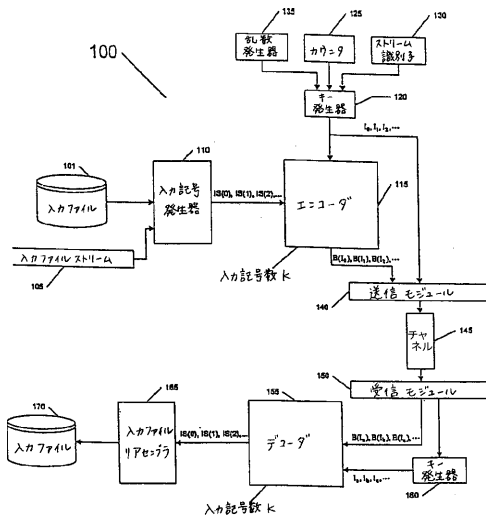
【図 2 6 c】図 2 6 c は、重み分布を実施するためのプログラムのソースコードリストである添付書類 A の例示である。

【図 2 6 d】図 2 6 d は、重み分布を実施するためのプログラムのソースコードリストである添付書類 A の例示である。

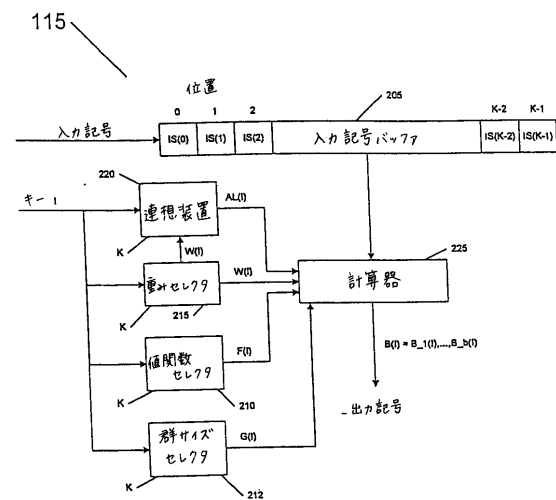
【図 2 6 e】図 2 6 e は、重み分布を実施するためのプログラムのソースコードリストである添付書類 A の例示である。

10

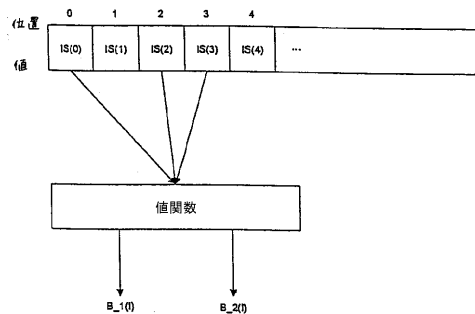
【図 1】



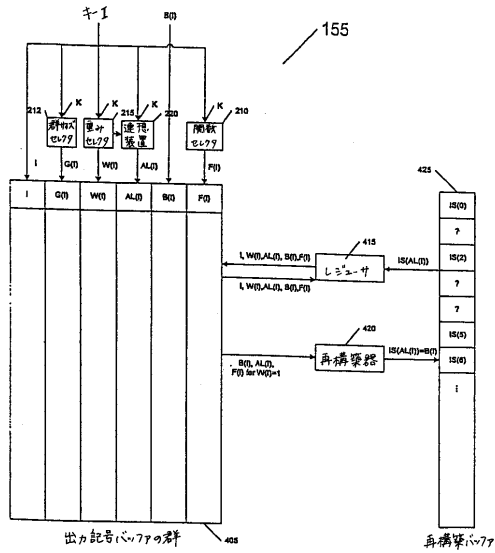
【図 2】



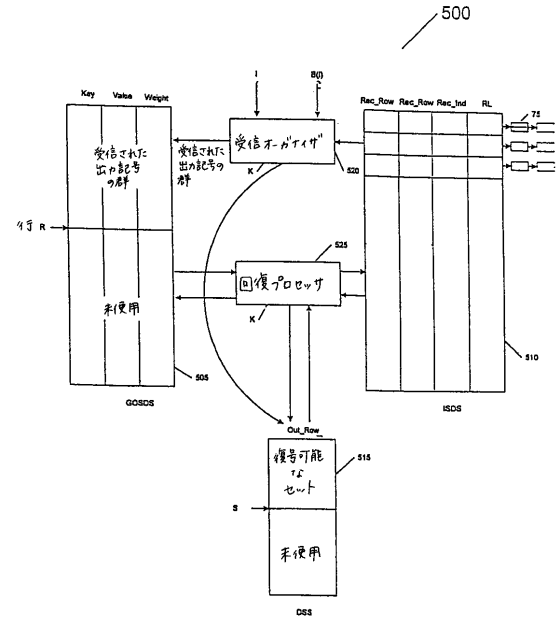
【図 3】



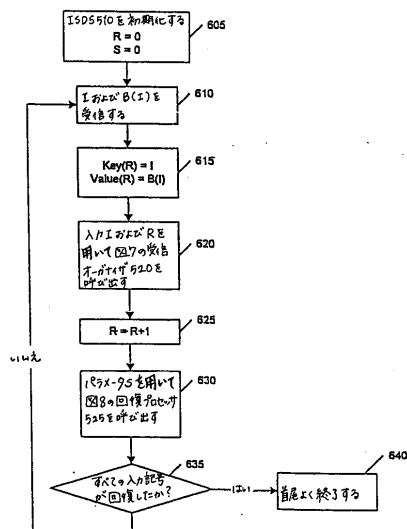
【図 4】



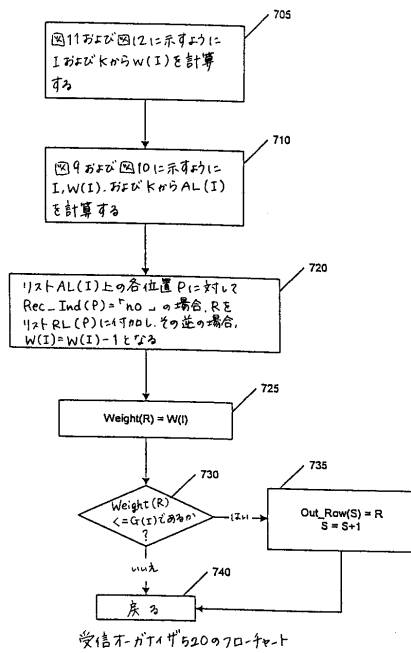
【図 5】



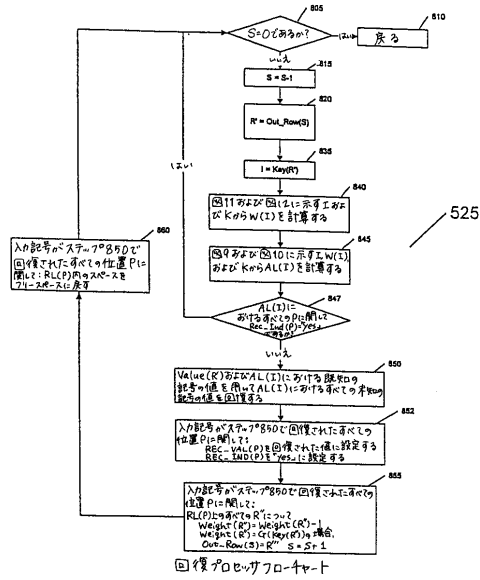
【図 6】



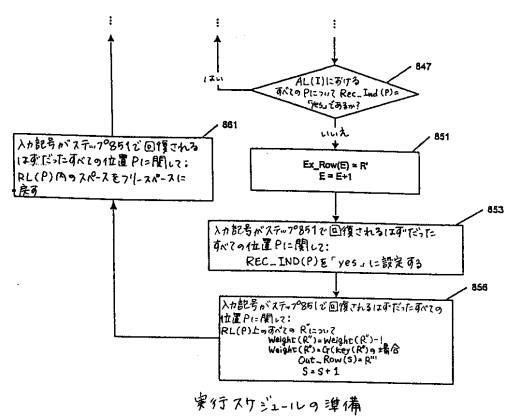
【図 7】



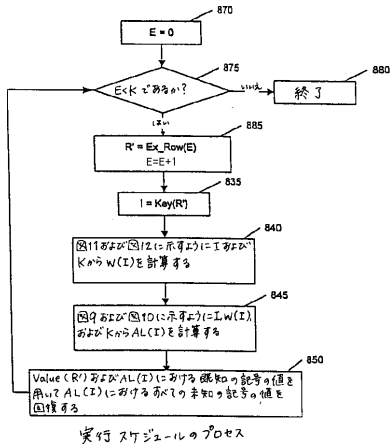
【図 8 a】



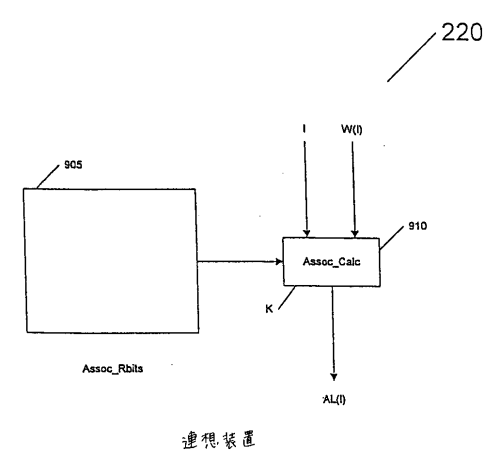
【図 8 b】



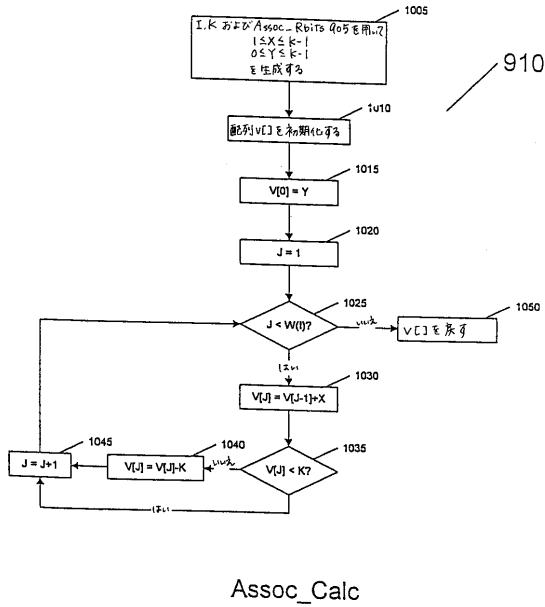
【図 8 c】



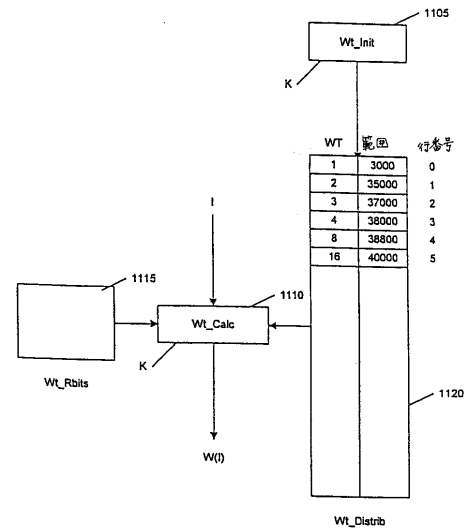
【図 9】



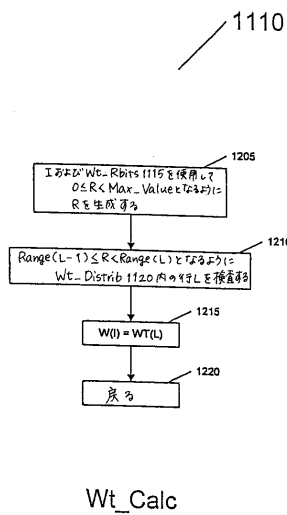
【図 10】



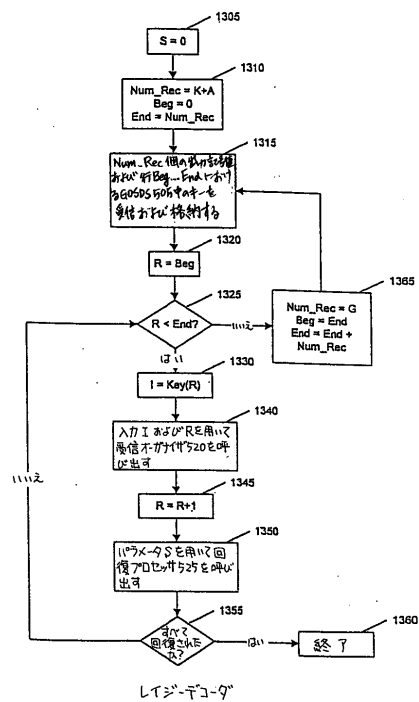
【図 11】



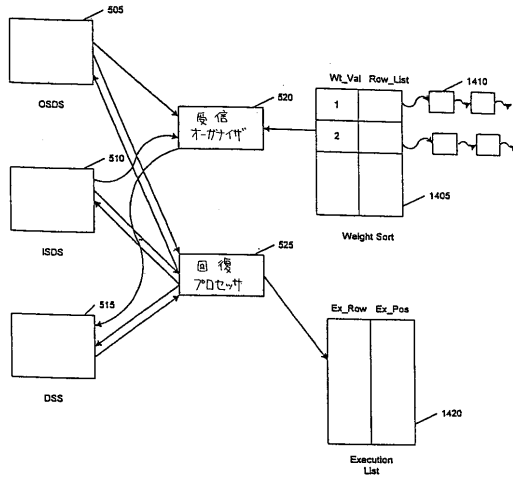
【図 12】



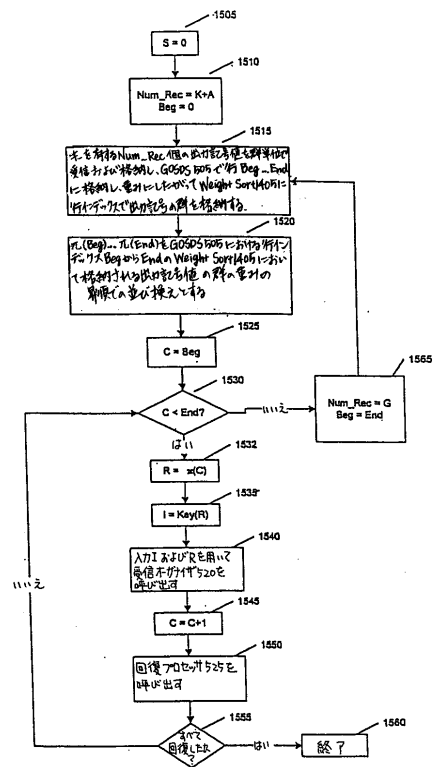
【図 13】



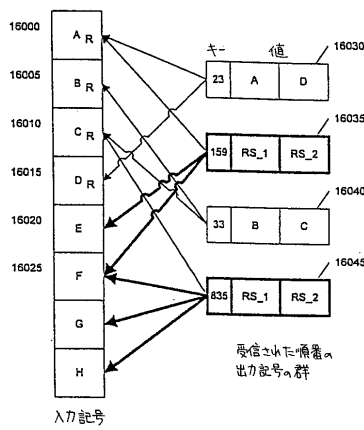
【図 14】



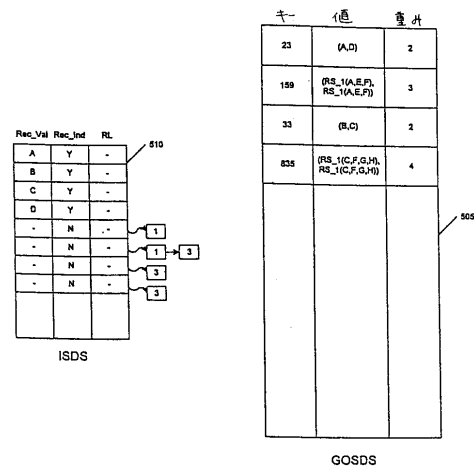
【図 15】



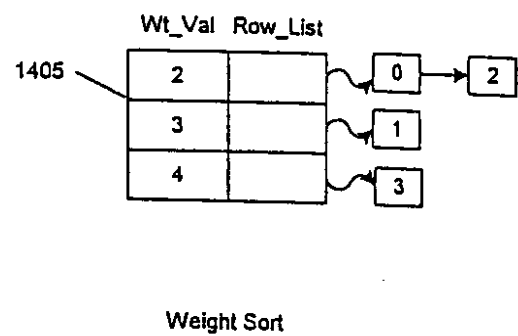
【図 16】



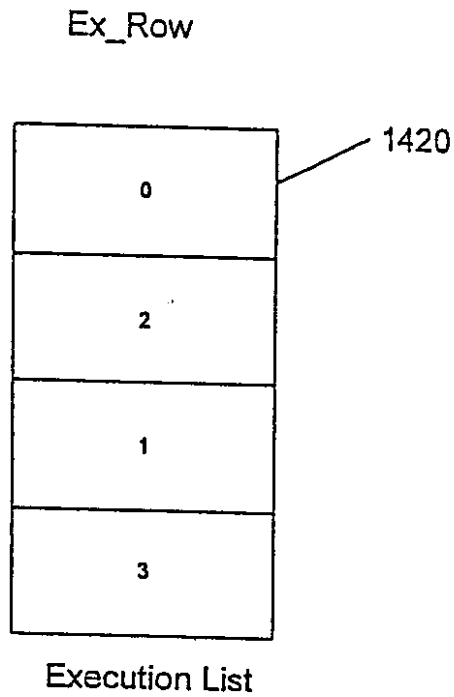
【図 17】



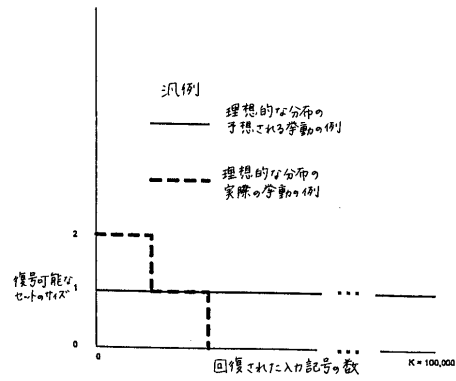
【図 18】



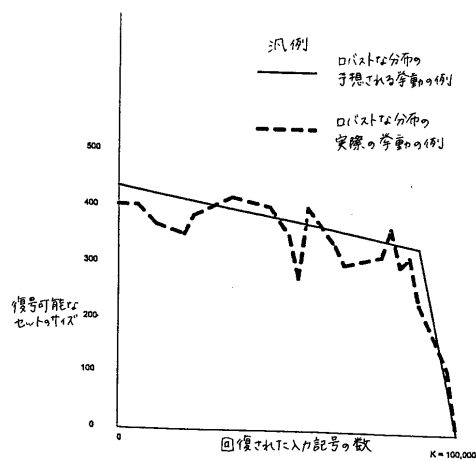
【図 19】



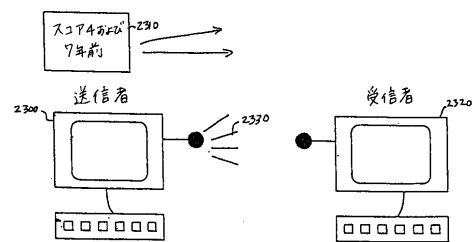
【図 20】



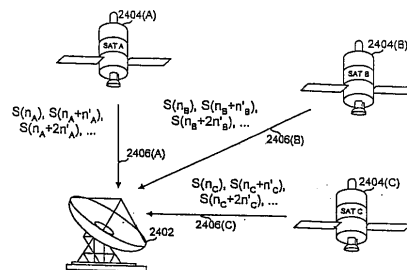
【図 21】



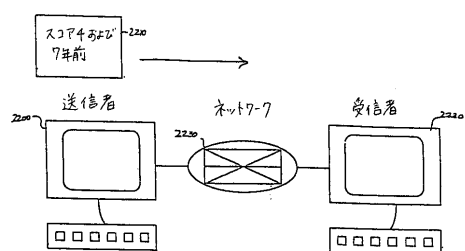
【図 23】



【図 24】



【図 22】





## 【図 26 d】

```

2. DfSolitonDistribution.h
//////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
// DfSolitonDistribution.h:
//
// CDFSolitonDistribution クラスに依存するインヘリタンス
//
// 注意: クラス CDFDominoKeyDecoder は、所与の出力記号キーの
// 近傍の重みおよびクラスを計算する。重みに対応する確率分布が二つのクラスを
// 参照し、選択可能なものに使用されるランダム整数を
// CDFRandomBits クラスを参照する
//
////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
#include "DFRandomBits.h"
// クラス CDFRandomBits;

//
// 分布の計算に使用される定数
//
////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
// 分布のビット単位の算術精度
// 分布密度は 2 のべき乗の乗数に尺度化される
const int knPrecision = 28;

// 「リップル-ゲートサイズ」とある R の計算に関連する定数
// この値は「ロバートソリトン分布」における予備されるリップル
// サイズである。R は値 kdrFactor * N の四乗根 + knRadd から
// 得られ、ここで N は入力記号数 (セグメント) である
const double kdrFactor = 2.1;
const int knRadd = 2;

// S は復号化開始時でのリップルの予備サイズであり、
// N 個の記号が受信される場合の重み 1 の出力記号の
// 予備される数と等しい
// S = kdrFactor * sqrt(N)
const double kdrFactor = 1.4;

// 分布の尾は、N/2, ..., N/R まで下った点での
// より高い値の密度である。
// 分布密度は、比例定数 kdrFactor と記号量に反比例する。
// 得られる分布は、上の上記の精度に
// 尺度化される

// 先例 1 を使用したが、それを 1.8 に増やし、最後の 1 の記号を
// 1 年間に長時間を要する「飛び地 (outliers)」を切り落とす。
// 次に 2.4 に増やし、最後の 4000 個の記号に与える (百万分の)
// 1 の飛び地に対応する。
const double kdrFactor = 2.4;

class CDFSolitonDistribution
{
    friend class CDFDominoKeyDecoder;

public:

```

## 【図 26 e】

```

////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
// 出力記号キーおよび入力記号の位置
//
// このクラスは (前に「インデックス」と呼んだ) 入力記号の位置を表す key-TKey
// および TPos を定義する。これらの key-TKey を必要とする他のクラスは以下の二つを
// 参照可能である
//
////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
// typedef 照り号 int TKey;
typedef unsigned _int32 TKey;
typedef int TPos;

CdfSolitonDistribution(
    int nSymbols,
    int nArity,
    int nRippleTargetSize = 0,
    int nStartRippleSize = 0);

virtual ~CdfSolitonDistribution();

inline double dAverageWeight() { return m_dAverageWeight; };
inline int nMaxWeight() { return m_nMaxWeight; };
inline int nParameterR() { return m_nRippleTargetSize; };
inline int nParameters() { return m_nStartRippleSize; };

// 確率分布は種類の配列を保持。
// この各種類は、出力記号の重み (前に
// 「重」と呼んだ) およびその重みに関連する
// 与えられる密度 (確率) に対応する

inline int nKinds() { return m_nKinds; };

private:
    double m_dAverageWeight;
    void ComputeAverageWeight();
    int m_nRippleTargetSize;
    int m_nStartRippleSize;

    // m_nKinds は確率分布を保持する
    // 配列のサイズである。可能なものは出力記号
    // の異なる重みの数である
    int m_nKinds;

    // 以下は不完全なロバートソリトン分布
    // からの種類数から生じる尾の
    // 分布による種類数である
    int m_nRobustKinds;
    int m_nTailKinds;

    int m_nModulus; // 精度に約 2

    int* m_anKindWeight; // 各の重みは種類に対応する
    int* m_anKindCumulativeDensity; // 種類の確率密度

    int m_nSymbols;
    int m_nArity;
    int m_nMaxWeight;
};

```

## フロントページの続き

(74)代理人 100075672  
弁理士 峰 隆司

(74)代理人 100095441  
弁理士 白根 俊郎

(74)代理人 100084618  
弁理士 村松 貞男

(74)代理人 100103034  
弁理士 野河 信久

(74)代理人 100119976  
弁理士 幸長 保次郎

(74)代理人 100153051  
弁理士 河野 直樹

(74)代理人 100140176  
弁理士 砂川 克

(74)代理人 100100952  
弁理士 風間 鉄也

(74)代理人 100101812  
弁理士 勝村 紘

(74)代理人 100070437  
弁理士 河井 将次

(74)代理人 100124394  
弁理士 佐藤 立志

(74)代理人 100112807  
弁理士 岡田 貴志

(74)代理人 100111073  
弁理士 堀内 美保子

(74)代理人 100134290  
弁理士 竹内 将訓

(74)代理人 100127144  
弁理士 市原 卓三

(74)代理人 100141933  
弁理士 山下 元

(74)代理人 100078282  
弁理士 山本 秀策

(72)発明者 マイケル ジー・ ルビー  
アメリカ合衆国 カリフォルニア 94708, パークレー, ミラー アベニュー 1133

審査官 渡辺 未央子

(56)参考文献 特開2001-036417(JP, A)  
特表2003-507985(JP, A)  
特許第3809957(JP, B2)  
特許第3976163(JP, B2)

(58)調査した分野(Int.Cl., DB名)

H03M 13/15  
H03M 13/09  
H04L 1/00