(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2014년06월26일
(11) 등록번호 10-1412036
(24) 등록일자 2014년06월19일

(51) 국제특허분류(Int. Cl.)
H03M 13/00 (2006.01) *H03M 7/30* (2006.01)
H04L 1/00 (2006.01)
(21) 출원번호 10-2013-0003294
(22) 출원일자 2013년01월11일
심사청구일자 2013년01월11일
(56) 선행기술조사문헌
KR1020100078405 A
Minhae Kwon 외1인, "Improved Approximate Decoding based on Position Information Matrix", ISCC 2012, 2012.07.1, pp.114-118.

(73) 특허권자
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
박형곤
경기 고양시 덕양구 화신로 106, 2416동 204호 (행신동, 햇빛마을24단지아파트)
권민혜
서울 은평구 연서로18길 33-2, 2층 (대조동)
(74) 대리인
김인철

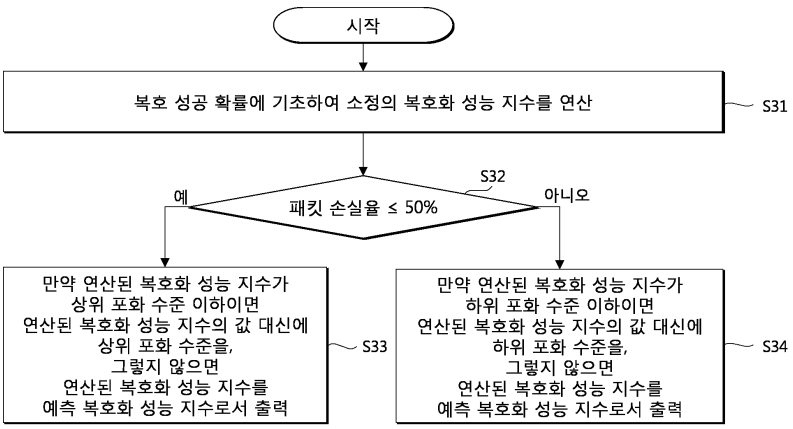
전체 청구항 수 : 총 5 항 심사관 : 성경아

(54) 발명의 명칭 **네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법**

(57) 요 약

본 발명의 실시예에 따른 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법은, GF(2^M) 내에서 연속하는 소스 데이터들 x_N 과 x_{N+1} 의 차이 $\Delta = x_N \oplus x_{N+1}$ 에 관한 각각의 후보 Δ_n 들 및 실제로 $\delta = x_N \oplus x_{N+1} = \Delta_n$ 인 Δ_n 의 위치 인덱스들에 기초하여 생성되는 위치 유사성 정보가 소스 데이터에 부가된 경우에, 위치 유사성 정보를 기초로 네트워크 코딩된 소스 데이터를 근사 복호화할 때의 복호 성공 확률에 따라 복호화 성능 지수를 연산하는 단계, 주어진 채널의 패킷 손실율이 50% 이하인 경우에, 만약 연산된 복호화 성능 지수가 상위 포화 수준 이하이면 연산된 복호화 성능 지수의 값 대신에 상위 포화 수준을, 그렇지 않으면 연산된 복호화 성능 지수를 예측 복호화 성능 지수로서 출력하는 단계 및 주어진 채널의 패킷 손실율이 50%를 초과하는 경우에, 만약 연산된 복호화 성능 지수가 하위 포화 수준 이하이면 연산된 복호화 성능 지수의 값 대신에 하위 포화 수준을, 그렇지 않으면 연산된 복호화 성능 지수를 예측 복호화 성능 지수로서 출력하는 단계를 포함할 수 있다.

대 표 도 - 도3



이 발명을 지원한 국가연구개발사업
과제고유번호 2010-0009717
부처명 교육과학기술부
연구사업명 기초연구사업-일반연구자지원사업-기본연구(유형2)
연구과제명 다양한네트워크환경에서강인한실시간멀티미디어전송전략
기 여 율 20/100
주관기관 이화여자대학교
연구기간 2010.05.01 ~ 2015.04.30
이 발명을 지원한 국가연구개발사업
과제고유번호 NIPA-2013-H0301-13-1002
부처명 미래창조과학부
연구사업명 대학IT연구센터육성지원사업
연구과제명 초고속무선네트워크기반u-Office서비스연구개발
기 여 율 40/100
주관기관 중앙대학교 산학협력단
연구기간 2013.01.01 ~ 2013.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 NIPA-2013-H0301-13-1008
부처명 미래창조과학부
연구사업명 대학IT연구센터육성지원사업
연구과제명 디지털 방송장비 연구(차세대디지털TV방송시스템(실감형/모바일/양방향)핵심기술개발)
기 여 율 40/100
주관기관 연세대학교 산학협력단
연구기간 2013.01.01 ~ 2013.12.31

특허청구의 범위

청구항 1

GF(2^M) 내에서 연속하는 소스 데이터들 x_N과 x_{N+1}의 차이 $\Delta = x_N \oplus x_{N+1}$에 관한 각각의 후보 Δ_n들 및 실제로 $\delta = x_N \oplus x_{N+1} = \Delta_n$인 Δ_n의 위치 인덱스들에 기초하여 생성되는 위치 유사성 정보가 소스 데이터에 부가된 경우에, 상기 위치 유사성 정보를 기초로 네트워크 코딩된 소스 데이터를 근사 복호화할 때의 복호화 성능 지수를 연산하는 단계;

주어진 채널의 패킷 손실율이 50% 이하인 경우에, 만약 연산된 복호화 성능 지수가 상위 포화 수준 이하이면 연산된 복호화 성능 지수의 값 대신에 상기 상위 포화 수준을, 그렇지 않으면 연산된 복호화 성능 지수를 예측 복호화 성능 지수로서 출력하는 단계; 및

주어진 채널의 패킷 손실율이 50%를 초과하는 경우에, 만약 연산된 복호화 성능 지수가 하위 포화 수준 이하이면 연산된 복호화 성능 지수의 값 대신에 상기 하위 포화 수준을, 그렇지 않으면 연산된 복호화 성능 지수를 예측 복호화 성능 지수로서 출력하는 단계를 포함하고,

상기 복호화 성능 지수는 복호 성공 확률이 높을수록 높은 값을 갖도록 연산되고, 패킷 손실율이 50% 이하인 경우에 상기 위치 유사성 정보의 크기가 줄어들면 상기 상위 포화 수준에 수렴하며, 패킷 손실율이 50% 초과인 경우에 상기 위치 유사성 정보의 크기가 줄어들면 상기 하위 포화 수준에 수렴하는 것을 특징으로 하는 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법.

청구항 2

청구항 1에 있어서, 상기 후보 Δ_n는

$$\Delta_n = \sum_{i=k}^{k+n-1} 2^i$$

여기서, 1≤n≤M-k이고 M은 GF의 특성값이며,

실제로 $\delta = x_N \oplus x_{N+1}$인 δ가 Δ_n일 확률은

$$\Pr(\delta = \Delta_n) = \frac{2^{(M-k-n)}}{2^{(M-k)} - 1} = \frac{2^{M-n}}{2^M - 2^k}$$

이고, 실제로 $\delta = x_N \oplus x_{N+1}$인 δ가 Δ_n일 확률이 큰 순서대로 선정되는 후보 Δ_n의 위치 인덱스들과, 소스 데이터에 위치 유사성 정보를 부가할 수 있는 크기에 기초하여, 상기 위치 유사성 정보가 생성되는 것을 특징으로 하는 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법.

청구항 3

청구항 2에 있어서, 상기 복호 성공 확률은

$$\Pr(\hat{\mathbf{s}} = \mathbf{s} | \Delta_1, \ldots, \Delta_n, N_l) = \left(\sum_{k=1}^{n} \frac{1}{2^k} \right)^{N_l}$$

로 근사되며, 손실 패킷이 N_l개인 조건에서 실수 공간에서의 복원 심볼 $\hat{\mathbf{s}}$이 n 개의 후보 Δ_n들을 이용하여 원본 심볼 $\mathbf{s}$과 동일하게 복원될 확률인 것을 특징으로 하는 네트워크 코딩된 소스 데이터의 위치 유사성 정보를

이용한 근사 복호화의 복호화 성능 예측 방법.

청구항 4

청구항 1에 있어서, 상기 복호화 성능 지수는 복호 성공 확률, 정규화된 오류율 또는 평균 제곱 오류율 중 어느 하나인 것을 특징으로 하는 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법.

청구항 5

컴퓨터에서 청구항 1 내지 청구항 4 중 어느 한 청구항에 따른 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법을 수행하도록 프로그래밍된 프로그램이 기록된 컴퓨터로 읽을 수 있는 기록 매체.

명 세 서

기 술 분 야

[0001] 본 발명은 네트워크 코딩에 관한 것으로, 더욱 상세하게는, 네트워크 코딩된 데이터의 근사 복호화 기술에 관한 것이다.

배 경 기 술

[0002] 무선 네트워크 기능을 가진 모바일 장치들이 일상적으로 사용됨에 따라, 무선 네트워크가 수많은 이기종의 장치들에 의해 지속적으로 변화하면서 동적으로 구성되고 있고, 이러한 동적인 네트워크 내에서 적시에 정보를 전송하는 기술의 중요성이 점차 높아지고 있다.

[0003] 네트워크 코딩(Network coding)은 경로와 소스가 다양한 동적 네트워크에서 효율적인 분산 전송 알고리즘을 구축할 수 있는 방법론으로서 점차 활용도를 넓혀 가고 있다. 기존의 라우팅 기반의 네트워크 스킴은 단순히 점대점으로 데이터를 그대로 송신(forwarding)하는 것이어서 중간에 네트워크 구성이 바뀌면 기존의 라우팅 경로가 끊기거나 속도가 저하되어 서비스 품질이 영향을 받고 라우팅 경로를 유지하기 위한 오버헤드가 적지 않다.

[0004] 이에 비하여, 네트워크 코딩은 네트워크 내의 인접한 노드들이 패킷들을 수신하고 결합하여 출력하는 동작을 반복함으로써 정해진 라우팅 경로 없이도 데이터들이 목적지까지 흘러 가도록 하는 것이므로 네트워크 코딩은 대역폭이나 전력 자원, 연산 자원을 효율적으로 사용할 수 있고, 네트워크의 변동에 강인함을 보인다. 소스와 경로가 다양할수록, 즉 동일한 정보를 가지고 있는 노드도 많고 요청하는 노드도 많을 경우에, 이러한 네트워크 코딩의 성능이 최대화될 수 있다.

[0005] 하지만 네트워크 코딩을 통해 전달되는 정보들은 데이터를 구성하는 패킷들이 원본 그대로 전달되는 것이 아니라 네트워크 노드가 적절한 연산을 통해 변조하기 때문에, 정해진 시간 내에 심볼들을 불완전하게 수신하였을 때에 데이터를 재구성할 수 없는 문제가 있다.

[0006] 특히 멀티미디어 데이터와 같이 시간 지연 민감성 데이터의 경우에는 허용 지연 시간 내에 패킷이 도착하지 않으면 종래의 디코딩 방법론들로는 데이터 복원 자체가 불가능하여 네트워크 코딩의 큰 한계로 지적되고 있다.

[0007] 이러한 문제점들에 대응하고자 제안된 근사 복호법은 네트워크 코딩된 데이터를 수신하였을 때에 수신된 데이터들의 상관 관계(correlation)를 이용하여 손실된 데이터를 근사적으로 복원할 수 있는 기법이다. 예를 들어 현실 세계에서 측정되는 온도나 전압은 시간적으로 근접한 측정 데이터 사이의 상관 관계가 높기 때문에, 데이터 수집 장치가 전송받은 데이터가 네트워크 복호화에 충분하지 못한 경우에도 원천 데이터의 상관성 정보를 이용하면 근사적으로 복원될 수 있다.

선행기술문헌

비특허문헌

[0008] (비특허문헌 0001) [1] Ho, T., M´edard, M., Shi, J., Effros, M. 및 Karger, D.R., "On randomized network coding," in Proc. Allerton Annual Conf. Commun., Control, and Comput., Oct. 2003

(비특허문헌 0002) [2] H. Park, N. Thomos, and P. Frossard, Approximate decoding approaches for

network coded correlated data, Signal Processing (Elsevier), vol. 93, no. 1, pp. 109~213, Jan. 2013

발명의 내용

해결하려는 과제

[0009] 본 발명이 해결하고자 하는 과제는 네트워크 코딩된 소스 데이터의 근사 복호화를 위한 위치 유사성 정보 크기의 최적화 방법을 제공하는 데에 있다.

과제의 해결 수단

[0010] 본 발명의 일 측면에 따른 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법은,

[0011] $GF(2^M)$ 내에서 연속하는 소스 데이터들 x_N과 x_{N+1}의 차이 $\Delta = x_N \oplus x_{N+1}$에 관한 각각의 후보 Δ_n들 및 실제로 $\delta = x_N \oplus x_{N+1} = \Delta_n$인 Δ_n의 위치 인덱스들에 기초하여 생성되는 위치 유사성 정보가 소스 데이터에 부가된 경우에, 상기 위치 유사성 정보를 기초로 네트워크 코딩된 소스 데이터를 근사 복호화할 때의 복호 성공 확률에 따라 복호화 성능 지수를 연산하는 단계;

[0012] 주어진 채널의 패킷 손실율이 50% 이하인 경우에, 만약 연산된 복호화 성능 지수가 상위 포화 수준 이하이면 연산된 복호화 성능 지수의 값 대신에 상기 상위 포화 수준을, 그렇지 않으면 연산된 복호화 성능 지수를 예측 복호화 성능 지수로서 출력하는 단계; 및

[0013] 주어진 채널의 패킷 손실율이 50%를 초과하는 경우에, 만약 연산된 복호화 성능 지수가 하위 포화 수준 이하이면 연산된 복호화 성능 지수의 값 대신에 상기 하위 포화 수준을, 그렇지 않으면 연산된 복호화 성능 지수를 예측 복호화 성능 지수로서 출력하는 단계를 포함할 수 있다.

[0014] 일 실시예에 따라, 상기 후보 Δ_n는

[0015] $$\Delta_n = \sum_{i=k}^{k+n-1} 2^i$$

[0016] 여기서, $1 \le n \le M-k$이고 M은 GF의 특성값이며,

[0017] 실제로 $\delta = x_N \oplus x_{N+1}$인 δ가 Δ_n일 확률은

[0018] $$\Pr(\delta = \Delta_n) = \frac{2^{(M-k-n)}}{2^{(M-k)} - 1} = \frac{2^{M-n}}{2^M - 2^k}$$

[0019] 일 수 있다.

[0020] 일 실시예에 따라, 상기 복호 성공 확률은

[0021] $$\Pr(\hat{\mathbf{s}} = \mathbf{s} | \Delta_1, \ldots, \Delta_n, N_l) = \left(\sum_{k=1}^{n} \frac{1}{2^k} \right)^{N_l}$$

[0022] 로 근사되며, 손실 패킷이 N_l개인 조건에서 실수 공간에서의 복원 심볼 $\hat{\mathbf{s}}$이 n 개의 후보 Δ_n들을 이용하여 원본 심볼 $\mathbf{s}$과 동일하게 복원될 확률일 수 있다.

[0023] 일 실시예에 따라, 상기 복호화 성능 지수는 복호 성공 확률, 정규화된 오류율 또는 평균 제곱 오류율 중 어느 하나일 수 있다..

발명의 효과

[0024] 본 발명의 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법에 따르면, 연속하는 소스 심볼들 또는 패킷들 사이의 위치 유사성 정보를 이용하여 복호화 성능을 예측함에 따라 수신 노드가 최적의 근사 복호를 수행할 수 있다.

[0025] 본 발명의 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법에 따르면, 채널의 패킷 손실도 또는 소스 노드의 인코딩 패킷 수 및 위치 유사성 정보의 크기에 따라 실제 복호화 성능을 통계로 추적하지 않고 예측할 수 있다.

도면의 간단한 설명

[0026] 도 1은 본 발명이 적용될 수 있는 분산된 데이터 전송 시스템을 개념적으로 표현한 다이어그램이다.

도 2는 본 발명의 일 실시예에 따른 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법에서 손실 패킷 수와 위치 유사성 정보의 크기에 따른 근사 복호화 성능을 예시한 그래프이다.

도 3은 본 발명의 일 실시예에 따른 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법을 예시한 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0027] 본문에 개시되어 있는 본 발명의 실시예들에 대해서, 특정한 구조적 내지 기능적 설명들은 단지 본 발명의 실시예를 설명하기 위한 목적으로 예시된 것으로, 본 발명의 실시예들은 다양한 형태로 실시될 수 있으며 본문에 설명된 실시예들에 한정되는 것으로 해석되어서는 아니 된다.

[0028] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다. 도면상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고 동일한 구성요소에 대해서 중복된 설명은 생략한다.

[0029] 아래의 설명은 크게 세 개의 섹션으로 나누어지는데, 섹션 1 및 2는 본 발명이 적용될 수 있는 분산된 데이터 전송 시스템에 관한 소개과 함께 근사 복호화법의 기본적 원리를 설명하며, 섹션 3은 본 발명에 따른 근사 복호화법을 설명한다.

[0030] 섹션 3에서 나타나는 벡터나 변수에 관한 표기 중 일부는 섹션 1 및 섹션 2에서 나타나는 것과 다를 수 있다는 점에 유의한다.

[0031] 또한 아래 설명에서 심볼, 패킷 및 데이터와 같은 용어들은 맥락 상 명백하게 다르게 해석되는 경우를 제외하면 서로 혼용하여 사용됨을 유의한다.

[0032] 1. 랜덤 선형 네트워크 코딩

[0033] 도 1은 본 발명이 적용될 수 있는 분산된 데이터 전송 시스템을 개념적으로 표현한 다이어그램이다. 소스 데이터가 네트워크 코딩 노드를 통하여 디코더(복호기)로 전달되는 과정이 나타나 있다. 이 시스템에서 전송되는 데이터는 서로 상관 관계를 갖는다. 상관 관계를 가지는 데이터는 센서 네트워크를 통해 측정된 데이터와 같은 외부 상관관계를 갖거나, 비디오 시퀀스 내의 이미지들과 같이 근접한 데이터 사이의 상관관계가 높은 데이터일 수 있다.

[0034] 이러한 상관관계 데이터의 전송에 대한 연구는 분산 시스템에서 널리 수행되어 왔다. 네트워크 코딩 기법은 네트워크를 통해 상관관계 데이터를 전송하는 기법으로 연구되었다.

[0035] 도 1에서 오버레이 애드 혹 네트워크(overlay ad hoc network)는 소스 노드들, 중간 노드들 및 수신 노드들로 구성되어 있고, 소스 데이터는 소스 노드에서 네트워크 코딩 연산을 담당하는 중간 노드들을 거쳐 수신 노드까지 전달된다. 중간 노드는 수신한 데이터를 그대로 전달할 수도 있지만, 다른 소스 노드들에서 수신한 데이터들을 결합(combining)하여 출력한다.

[0036] 다양한 네트워크 코딩 기법 중에서 본 발명에서 사용하는 랜덤 선형 네트워크 코딩(RLNC: Random Linear Network Coding) 기법은 Ho, T., M´edard, M., Shi, J., Effros, M. 및 Karger, D.R., "On randomized

network coding," in Proc. Allerton Annual Conf. Commun., Control, and Comput., Oct. 2003을 참조한다.

[0037] 소스 집합 $\mathbf{S_n}$ 내의 L 개의 소스 심볼들은 $s_N^{(1)}, s_N^{(2)}, \ldots, s_N^{(L)}$ 라고 표시되며, 각각 이산(discrete)이면서 상관되어(correlated) 있고, 1≤i≤L일 때에 $s_N^{(i)} \in \mathcal{S}(\subset \mathbb{R})$ 이다. $\mathcal{S}$ 는 $s_N^{(i)}$ 의 알파벳 집합(alphabet set)이고 $|\mathcal{S}|$ 는 $\mathcal{S}$ 의 크기를 의미한다. 네트워크 코딩 연산은 갈로와 필드(GF: Galois Field) 내에서 수행되기 때문에, 각각의 $s_N^{(i)}$ 은 GF 내의 원소로 고려되어야 한다. $s_N^{(i)}$ 가 실수 필드($\mathbb{R}$, field of real numbers)에 있는 수인지 아니면 GF에 있는 수인지 여부를 명확하게 특정하기 위해서, 수학식 1과 같은 식별 함수(identity function)을 정의할 수 있다.

수학식 1

[0038]
$$\begin{cases} 1_{\mathbb{R}G} : \mathbb{R} \to \mathrm{GF}, & 1_{\mathbb{R}G}\left(s_N^{(i)}\right) = x_N^{(i)} \\ 1_{G\mathbb{R}} : \mathrm{GF} \to \mathbb{R}, & 1_{G\mathbb{R}}\left(x_N^{(i)}\right) = s_N^{(i)} \end{cases}$$

[0039] 이는 $x_N^{(i)}$ 가 $s_N^{(i)}$ 을 의미하는 GF 내의 원소임을 의미한다. $x_N^{(i)} \in \mathcal{X}$ 인 경우에, $\mathbf{x}_N = \left[x_N^{(1)}, \ldots, x_N^{(L)}\right]^T$ 는 L 개의 요소를 가지는 N 번째 소스 데이터 집합이다. $\mathcal{X}$ 는 $x_N^{(i)}$ 의 알파벳 세트를 의미한다. 또한 본 명세서에서 행렬 또는 벡터의 위 첨자로 쓰이는 T는 행렬 또는 벡터의 전치(transpose)를 의미하며, 실수값 T 또는 아래 첨자로 쓰이는 T와 구별됨을 유의한다.

[0040] GF 내의 연산 기호를 가지고 서술하는 수식에서는 모든 연산은 GF 내에서 일어나는 것을 암시하므로, $1_{\mathbb{R}G}$ 기호는 생략된다. $\oplus$ 와 $\otimes$ 연산자들은 각각 GF 내에서 정의되는 합산 연산과 승산 연산을 의미한다.

[0041] RLNC를 이용하는 중간 노드 k는 다음 수학식 2와 같이 패킷을 생성하여 전송한다.

수학식 2

[0042]
$$y(k) = \sum_{t=1}^{T} \bigoplus \left\{ c_t(k) \otimes x_t^{(i)} \right\}$$
$$\triangleq \left\{ c_1(k) \otimes x_1^{(i)} \right\} \oplus \left\{ c_2(k) \otimes x_2^{(i)} \right\} \oplus \cdots \oplus \left\{ c_T(k) \otimes x_T^{(i)} \right\}$$

[0043] 수학식 2는 $x_t^{(i)}$ 와 코딩 계수 $c_t(k)$ 의 GF에서의 선형 결합(linear combination in GF)을 의미한다. i는 1≤i≤L인 임의의 정수이고, T는 하나의 패킷 내에 결합된 데이터의 개수이다. $y(k)$ 와 $c_t(k)$ 는 항상 GF 내의 원소로 간주될 수 있다. 코딩 계수들은 크기 2^M 을 가지고 GF로부터 균일하고 랜덤하게 선정되는데, 이를 $\mathrm{GF}(2^M)$ 라고 한다. 이는 GF 크기가 M에 의해 결정되며, $c_t(k) \in \mathrm{GF}\left(2^M\right)$ 임을 의미한다.

[0044] 본 발명에서는 특히, 특성값 2를 가지는 GF, 즉 $\mathrm{GF}(2^M)$ 에서 합산 연산은 XOR 연산으로 수행될 수 있다. GF의 크기는 소스 심볼에 관하여 수행될 수 있는 코딩 연산들의 집합을 결정한다. 따라서 입력 알파벳의 크기가

$|\mathcal{S}| \le 2^M$라고 가정할 수 있다. 만약 $|\mathcal{S}| > 2^M$인 경우에는 입력 집합은 예를 들어 소스 결합(source binning) 또는 양자화 등을 이용하여 축소되고 그럼으로써 입력 집합은 GF 크기, 즉 2^M을 초과하지 않게 된다.

[0045] 각각의 노드에서 생성되는 패킷들은 이웃하는 노드들로 전송되며 최종적으로 수신 노드들로 향한다. 만약 수신 노드의 디코더에서 K 개의 이노배티브(innovative) 즉, 선형 비종속적(linear independent)인 심볼 또는 패킷들, $y(1), \ldots, y(K)$이 수신되어 연산에 쓰일 수 있다면, 선형 시스템 $\mathbf{y} = \mathbf{C} \odot \mathbf{x}$는 다음 수학식 3과 같이 표현될 수 있다.

수학식 3

[0046]
$$\begin{bmatrix} y(1) \\ \vdots \\ y(K) \end{bmatrix} = [\mathbf{c}_1 \ \cdots \ \mathbf{c}_T] \odot \begin{bmatrix} x_1^{(i)} \\ \vdots \\ x_T^{(i)} \end{bmatrix} \triangleq \sum_{t=1}^{T} \bigoplus \{\mathbf{c}_t \otimes x_t^{(i)}\}$$

[0047] 여기서 $\odot$ 연산자는 유한 필드(finite field)에서 행렬들 사이의 승산 연산을 의미한다.

[0048] K×T 행렬 $\mathbf{C}$는 코딩 계수 행렬(coding coefficient matrix)이라고 하는데, 컬럼 벡터들 $\mathbf{c}_t = [c_t(1), c_t(2), \cdots, c_t(K)]^T$로 구성되며, 수신된 심볼들의 유사도 모델에 기초하여 생성된다. 코딩 계수 행렬 $\mathbf{C}$를 유사도 모델에 기초하여 생성하는 방법은 네트워크 코딩의 복호화 기법에 숙련된 자들에게 주지된 방법이므로 설명을 생략한다.

[0049] 만약 이러한 코딩 계수 행렬 $\mathbf{C}$가 풀-랭크(full-rank), 즉 K=T인 경우에, 수학식 3의 선형 시스템의 해(solution)인 $\hat{\mathbf{x}} = [\hat{x}_1^{(i)}, \ldots, \hat{x}_T^{(i)}]^T$는, 코딩 계수 행렬 $\mathbf{C}$의 역행렬 $\mathbf{C}^{-1}$가 고유하게 결정되고 또한 이에 상응하여 $\hat{\mathbf{s}} = \mathbf{1}_{G\mathbb{R}}(\hat{\mathbf{x}})$이기 때문에, 다음 수학식 4와 같이 고유하게 결정될 수 있다.

수학식 4

[0050]
$$\hat{\mathbf{x}}_N = [\hat{x}_1^{(i)}, \ldots, \hat{x}_T^{(i)}]^T = \mathbf{C}^{-1} \odot \mathbf{y}$$

[0051] 코딩 계수 행렬의 역행렬은 GF 상에서 가우시안 소거법(Gaussian elimination)을 수행하는 것과 같은 주지의 방법을 통해 쉽게 얻을 수 있다.

[0052] 문제는, 만약 고유한 역행렬 $\mathbf{C}^{-1}$을 결정하기에 충분하지 않은 수의 심볼들만 수신된 경우에, 다시 말해, K<T인 경우에, 코딩 계수 행렬 $\mathbf{C}$가 풀-랭크가 아니기 때문에 수학식 3의 선형 시스템의 해 $\hat{\mathbf{x}} = [\hat{x}_1^{(i)}, \ldots, \hat{x}_T^{(i)}]^T$는 유일하게 결정되지 못할 뿐 아니라, 무한한 개수로 얻어질 수 있다.

[0053] 근사 복호화법은 이러한 문제를 해결하기 위해 도입된 것으로, 근사적으로 결정된 소스 심볼들이 불충분하게 수신된 데이터를 가지고 재구축된다.

[0054] 2. 근사 복호화법

[0055] 복호기(decoder)는 심볼 집합 y를 수신하면 소스 데이터를 복원할 수 있다. 만약 전송 절차 중에 패킷 지연이

나 손실과 같은 외부적 요인이 있거나, 또는 무작위로 선정되는 코딩 계수 행렬 $\mathbf{C}$이 특이 행렬(singular matrix)인 경우에는, 코딩 계수 행렬 $\mathbf{C}$의 역행렬 $\mathbf{C}^{-1}$은 얻어질 수 없고 수학식 3의 선형 시스템에서 해가 유일하지 않을 수 있다.

[0056] 이러한 문제를 해결하기 위해 근사 복호화법은 추가적인 제한 조건들(additional constraints)을 이용하여 코딩 계수 행렬을 풀-랭크로 만들 수 있다.

[0057] 입력된 데이터의 상관 관계가 복호화 절차에서 추가되는 추가 제한 조건들 $\mathbf{D}$와 $\boldsymbol{\nu}$를 구축하는 데에 이용될 수 있다. 추가 제한 조건들 $\mathbf{D}$ 및 $\boldsymbol{\nu}$의 모든 요소들(elements) 역시 GF 내의 값이다.

[0058] 추가 제한 조건들 $\mathbf{D}$ 및 $\boldsymbol{\nu}$을 추가함으로써 선형 시스템은 유일한 해를 가질 수 있게 되고 유일한 해 $\hat{\mathbf{x}}$는 다음 수학식 5와 같이 표현될 수 있다.

수학식 5

[0059]
$$\hat{\mathbf{x}} = \begin{bmatrix} \mathbf{C} \\ \mathbf{D} \end{bmatrix}^{-1} \odot \begin{bmatrix} \mathbf{y} \\ \boldsymbol{\nu} \end{bmatrix}$$

[0060] 추가 제한 조건들 $\mathbf{D}$ 및 $\boldsymbol{\nu}$ 내의 계수들은 소스의 모델들에 기초하여 결정될 수 있다. $\hat{\mathbf{x}}$가 결정되면, 원본 데이터의 근사 심볼 $\hat{\mathbf{s}}$은 수학식 1의 식별 함수들에 의해 $\hat{\mathbf{s}} = 1_{GR}(\hat{\mathbf{x}})$과 같이 얻어질 수 있다.

[0061] 좀더 상세한 내용은 H. Park, N. Thomos, and P. Frossard, Approximate decoding approaches for network coded correlated data, Signal Processing (Elsevier), vol. 93, no. 1, pp. 109~213, Jan. 2013 논문을 참조할 수 있다.

[0062] (T-K)×T 행렬 $\mathbf{D}$는 유사도 모델에 기초하여 생성될 수 있는데, 예를 들어 가장 유사한 데이터 $x_t^{(i)}$ 및 $x_{t'}^{(i)}$의 위치에 상응하는 자리에 값 "1" 및 "1"을 가지는 두 개의 요소들을 제외하고 각 행이 0으로 구성되도록 생성된다. 또한 $\boldsymbol{\nu}$은 T-K 길이를 가지는 벡터로서 결정되는데, 예를 들어 모든 값이 0일 수 있다 ($\boldsymbol{\nu} = \mathbf{0}_{T-K}$).

[0063] 이렇게 하여 원본 데이터의 근사 심볼 $\hat{\mathbf{s}}$은 다음의 수학식 6과 같이 얻어질 수 있다.

수학식 6

[0064]
$$\hat{\mathbf{s}} = 1_{GR}(\hat{\mathbf{x}}) = 1_{GR}\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{D} \end{bmatrix}^{-1} \odot \begin{bmatrix} \mathbf{y} \\ \mathbf{0}_{(T-K)} \end{bmatrix}\right)$$

[0065] 이를 통해 복호화기는 일부 심볼들이 완전한 복호화에 불충분하더라도 원래의 심볼들을 근사적으로 복원할 수 있다.

[0066] 이러한 기존의 알고리즘의 핵심 아이디어는 소스의 상관 관계에 기초하여 잉여 방정식을 추가하여 수학식 6에서 $\begin{bmatrix}\mathbf{C}^T & \mathbf{D}^T\end{bmatrix}^T$를 가역적으로(invertible) 만든다는 점이다. 이는 네트워크 코딩된 데이터에 관하여 데이터

복원 방법론의 새로운 패러다임을 제시한다.

[0067] 하지만 이렇게 기존에 제안된 근사 복호화 방법론들은 추가 제한 조건 행렬 $\mathbf{D}$와 $\boldsymbol{\nu}$가 경험론적 방법론(heuristic approach)에 의존하는 경우가 대부분이어서 소스 유사성을 제대로 재현하지 못하였고, 이는 제한된 성능으로 이어졌다.

[0068] 3. 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화법과 성능 예측 기법

[0069] 3.1 소스 데이터의 위치 유사성 정보

[0070] 한편, $\boldsymbol{\Delta} = \boldsymbol{x}_{\mathbf{N}} \oplus \boldsymbol{x}_{\mathbf{N+1}}$ 인 $\boldsymbol{\Delta}$ 가 실수 공간에서는 하나의 값을 가지지만, GF(2^M) 내에서는 복수의 후보 Δ들, 즉 Δ_1, Δ_2, Δ_3, ..., Δ_n을 가질 수 있다는 점은 주지의 사실이다. 종래에는 이러한 특성이 근사 복호화법에 이용되지 않았지만, 본 발명과 동일 출원인이 2012년 6월 29일자로 출원한 한국특허출원 제2012-0071282호 "위치정보행렬을 활용한 네트워크 코딩된 데이터의 근사 복호화 방법"에서는 이러한 복수의 후보 값들과 그 위치 인덱스를 이용하는 근사 복호화 방법을 개시하며, 상술한 한국특허출원 제2012-0071282호는 본 발명의 내용의 일부로서 병합된다.

[0071] 상술한 한국특허출원 제2012-0071282호에서, 각각의 후보 Δ값 및 직전의 심볼에 대해 해당 후보 Δ값을 가지는 심볼의 위치에 관한 위치 인덱스들에 관한 정보를 후보 Δ 순으로 나열한 "위치정보행렬" 즉 PIM(Position Information Matrix)이라는 명명하고, 소스 데이터와 함께 수신 노드로 전송함으로써, 수신 노드에서 누락된 심볼들에 대한 근사 복호화를 수행할 때에 추가 제한 조건 $\boldsymbol{\nu}$을 생성하는 데에 PIM 정보들을 활용할 수 있게 한 발명이 개시된다.

[0072] GF(2^M) 내에서 만약 $\Delta=2^k$ ($0 \le k \le M$)라면 $\boldsymbol{\Delta}$ 은 M-k 개의 후보들을 가지는데, $\boldsymbol{\Delta}$에 속하는 n 번째 후보 Δ_n는 다음의 수학식 7과 같이 표현되고, 후보 Δ_n가 실제로 $\boldsymbol{\delta} = \boldsymbol{x}_{\mathbf{N}} \oplus \boldsymbol{x}_{\mathbf{N+1}}$ 인 $\boldsymbol{\delta}$와 일치할 확률은 수학식 8과 같이 표현된다.

수학식 7

[0073]
$$\Delta_n = \sum_{i=k}^{k+n-1} 2^i$$

수학식 8

[0074]
$$\Pr(\delta = \Delta_n) = \frac{2^{(M-k-n)}}{2^{(M-k)} - 1} = \frac{2^{M-n}}{2^M - 2^k}$$

[0075] 여기서, $1 \le n \le M-k$이다. 수학식 8의 확률은 GF(2^M) 내에서 전체 $2^{(M-K)}-1$ 회의 후보 Δ들의 가능성 중에 특정 Δ_n가 등장하는 횟수가 $2^{(M-k-n)}$임을 의미한다. 실제의 $\boldsymbol{\delta}$ 는 후보 Δ들 중 어느 하나이게 되므로, 수학식 8과 같은 확률식이 도출된다.

[0076] 예를 들어, $\boldsymbol{\delta}=\Delta_1=2^k$일 경우가 가장 확률이 높은데, 만약 소스 데이터 중에서 각각의 직전 심볼에 관하여 $\boldsymbol{\delta}=\Delta_1=2^k$로 연산되는 일부 심볼들의 위치 인덱스들에 관한 정보를 수신 노드가 알고 있다면, 소스 데이터 집합 내의 심볼들은 선형 상관 관계가 있으므로, 일부 심볼이 전송 중에 누락되더라도, 알고 있는 위치 인덱스를 이용하여 추가 제한 조건 $\boldsymbol{\nu}$를 생성한다면 좀더 높은 확률로 근사 복호할 수 있다는 것이 상술한 한국특허출원 제2012-

0071282호의 기술적 사상 중 하나라고 할 수 있다. 그러한 위치 인덱스는 소스 노드가 생성하여 수신 노드에 부가 정보로서 전달한다.

[0077] 나아가, 만약 Δ_1 다음으로 확률이 높은 Δ_2로 연산되는 심볼들의 위치 인덱스도 함께 부가하여 전송한다면, 나아가, 추가적으로 최대 M-k 개의 후보들 Δ_n에 관한 심볼들의 위치 인덱스도 부가하여 전송한다면, 수신 노드는 누락된 심볼들을 더욱 정확하게 근사 복호할 수 있는 것이다.

[0078] 본 발명에서는 이러한 Δ_1에 관한 정보와, 연속하는 심볼들 사이의 차이가 Δ_n로 나타나는 심볼들에 관한 위치 인덱스들의 정보를 확률 순서대로 나열하여 PSI(Position Similarity Information), 즉 위치 유사성 정보라고 한다.

[0079] 3.2 위치 유사성 정보를 이용한 근사 복호화와 성능 예측

[0080] 상술한 한국특허출원 제2012-0071282호에서 PIM이라 명명한 개념이 본 명세서에서는 위치 유사성 정보라고 호칭하기로 한다.

[0081] PSI 정보는 n의 크기에 따라 커지거나 작아질 수 있지만, 원래 소스 데이터에 부가되어 전송되는 것이므로 대역폭을 차지하게 된다. 따라서, PSI의 크기를 증가시키면 근사 복호가 좀더 정확하게 되는 대신에 전체 대역폭 중 데이터가 차지할 수 있는 대역폭을 감소시킨다. 반면에 데이터의 PSI의 크기를 축소시키면 전체 대역폭 중 데이터가 차지할 수 있는 대역폭이 늘어나는 대신에 근사 복호가 다소 부정확하게 된다.

[0082] n의 크기를 결정하는 요소를 알기 위해, 한 번에 인코딩 수만큼 결합되는 심볼들을 전송 중에 N_l 개의 심볼들이 소실된다고 가정하면, 수신 노드에서는 손실된 심볼의 갯수만큼의 길이를 가지는 벡터, 추가 제한 조건 $\boldsymbol{\nu}$를 PSI로부터 생성할 수 있다.

[0083] n이 1이면 수학식 8에 의해 $\boldsymbol{\Delta}$ 내에 약 1/2의 확률로 내에 Δ_1가 나타날 것이고, n이 2이면 약 $1/2^2$의 확률로 Δ_2가 $\boldsymbol{\Delta}$ 내에 나타난다. 따라서, n 개의 후보 Δ_1, Δ_2, Δ_3, ..., Δ_n들을 PSI에 포함시켰다면 하나의 손실 심볼에 대한 길이 1의 $\boldsymbol{\Delta}$ 내에 Δ_1, Δ_2, Δ_3, ..., Δ_n들이 모두 나타날 확률은 각각의 확률들을 모두 더한 값이다.

[0084] 만약 N_l개의 손실 심볼이라면, 다음 수학식 9와 같이 근사적으로 표현될 수 있다.

수학식 9

[0085]
$$\Pr(\hat{\mathbf{s}} = \mathbf{s} | \Delta_1, \ldots, \Delta_n, N_l) = \left(\sum_{k=1}^{n} \frac{1}{2^k} \right)^{N_l}$$

[0086] 이는 손실 심볼 N_l개인 조건에서 복원 심볼 $\hat{\mathbf{s}}$이 n 개의 후보 Δ들을 이용하여 정확히 원본 심볼 $\mathbf{s}$과 동일하게 복원될 확률을 의미하며, n과 k는 소스 노드에서 결정되므로, 수학식 9의 복원 확률, 나아가 수신 노드의 데이터 복원 성능은 손실 심볼들의 개수 N_l에 지배적으로 좌우된다고 할 수 있다.

[0087] 수학식 9에서 근사적으로 예측한 바에 따라, 손실 심볼들의 개수가 어느 정도로 성능을 결정하는지 보기 위해 도 2를 참조하면, 도 2는 본 발명의 일 실시예에 따른 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법에서 손실 패킷 수와 위치 유사성 정보의 크기에 따른 근사 복호화 성능을 예시한 그래프이다.

[0088] 여기서, 네트워크 코딩되는 데이터, 즉 패킷들의 수는 9 개이며, 채널의 패킷 손실율은 패킷 손실이 전체 패킷 9 개 중 하나인 경우부터 9 개중 8 개인 경우까지 다양한 상황을 나타낸다. 다만 $\boldsymbol{\Delta}$의 크기는 8이며 GF(2^{10})

로 가정한다. 복호화 성능은 평균 제곱 오류율(MSE)의 역수로 측정하였다.

[0089] 손실 패킷이 1 이상 4 이하인 경우에는, n=1인 Δ_1외에 PSI 부가 정보가 전혀 없을 때 즉 PSI 오버헤드 0%인 때에는 1/MSE 값들은 모두 상위 포화 수준(Sat_u) 근처에 있다.

[0090] 또한 손실 패킷이 5 이상 8 이하인 경우에는, n=1인 Δ_1외에 PSI 부가 정보가 전혀 없을 때 즉 PSI 오버헤드 0%인 때에는 1/MSE 값들은 모두 하위 포화 수준(Sat_l) 근처에 있다.

[0091] 한편, PSI 부가 정보가 n=2인 Δ_2까지 부가되면, 전송되어야 하는 데이터의 크기는 원본 데이터의 크기보다 약 20% 만큼 늘어나는데, 그 대신에 1/MSE 값들은 손실 패킷의 수가 1인 사례와 5인 사례에서 다소 개선된다. 하지만, 손실 패킷 수가 2 이상 4 이하인 사례들에서는 상위 포화 수준에, 6 이상 8이하인 사례들에서는 하위 포화 수준에 머무른다.

[0092] PSI 부가 정보가 n=3인 Δ_3까지 부가되면, 전송되어야 하는 데이터의 크기는 원본 데이터의 크기보다 약 30%만큼 늘어나는데, 그 대신에 1/MSE 값들은 손실 패킷의 수가 1 및 5인 사례에서 더욱 개선되었을 뿐 아니라, 손실 패킷 수가 각각 2 및 3인 사례와 6 및 7인 사례에서도 개선되기 시작한다. 하지만, 여전히 손실 패킷 수가 4인 사례에서는 상위 포화 수준에 머무르고, 손실 패킷 수가 8인 사례에서는 하위 포화 수준에 머무른다.

[0093] PSI 부가 정보가 n=4인 Δ_4까지 부가되면, 전송되어야 하는 데이터의 크기는 원본 데이터의 크기보다 약 36%만큼 늘어나는데, 그 대신에 1/MSE 값들은 상황에 따라 1/10 이상 줄어들 수 있다. 또한 모든 사례들에서 1/MSE 값들이 개선되었고, 특히 손실 패킷 수가 4인 사례와 손실 패킷 수가 8인 사례에서도 각각 상위 포화 수준과 하위 포화 수준을 이탈하여 급격한 성능 개선을 나타낸다.

[0094] 부가 정보가 n=5 이상으로 부가되면, 전송되어야 하는 데이터의 크기는 약 40% 정도로 늘어나지만, 모든 사례들에서 1/MSE 값들이 대단히 개선되는 점을 알 수 있다. 특히 손실 패킷 수가 1 이상 4 이하인 경우에 나타나는 1/MSE 값은 복호 성공 확률로 환산하면 약 95 ~ 99% 범위라고 할 수 있다.

[0095] 도 2를 관찰하면, 패킷 손실율이 50%보다 낮은 경우에 손실 패킷 수에 따른 복호화 성능 지수가 수렴하는 수준과 패킷 손실율이 50%보다 높은 경우에 손실 패킷 수에 따른 복호화 성능 지수가 수렴하는 수준이 다르다는 점이 다르다는 점이 특이하다. 또한 PSI 부가 정보의 부가에 따라, 패킷 손실율 0% 이상 50% 이하인 사례들과 패킷 손실율이 50% 초과 100% 미만인 사례들에서 각각 유사한 행동을 나타낸다는 점도 특이하다.

[0096] 어느 경우이든 PSI 부가 정보의 크기를 늘리면, 즉 전체 대역폭에서 소스 데이터를 위한 대역폭을 희생하면, 근사 복호화 성능 지수는 개선된다. 복호화 성능 지수는 이러한 복호 성공 확률 뿐 아니라 정규화 오류율나 평균 제곱 오류율(MSE) 또는 그 역수로 측량될 수 있는데, 목표로 하는 복호화 성능 지수와 대역폭 사이에는 트레이드오프 관계가 있다.

[0097] 따라서, 네트워크 코딩의 복호화 성능 지수를 예측할 수 있으면 네트워크 자원을 좀더 효율적으로 활용할 수 있는 방안이 도출될 수 있다.

[0098] 도 3은 본 발명의 일 실시예에 따른 네트워크 코딩된 소스 데이터의 위치 유사성 정보를 이용한 근사 복호화의 복호화 성능 예측 방법을 예시한 순서도이다.

[0099] 도 3을 참조하면, $GF(2^M)$ 내에서 연속하는 소스 데이터들 $\boldsymbol{x_N}$과 $\boldsymbol{x_{N+1}}$의 차이 $\boldsymbol{\Delta = x_N \oplus x_{N+1}}$에 관한 각각의 후보 $\boldsymbol{\Delta_n}$들 및 실제로 $\boldsymbol{\delta = x_N \oplus x_{N+1} = \Delta_n}$인 $\boldsymbol{\Delta_n}$의 위치 인덱스들에 기초하여 생성되는 위치 유사성 정보(PSI)가 소스 데이터에 부가된 경우를 전제로 하여, 단계(S31)에서, 수학식 9의 복호 성공 확률에 기초하여 소정의 복호화 성능 지수를 연산한다. 상술하였듯이, 복호화 성능 지수는 복호 성공 확률 자체일 수도 있고, 정규화된 오류율이나 평균 제곱 오류율일 수도 있다.

[0100] 단계(S32)에서, 주어진 채널의 패킷 손실율 γ이 50% 이하인지 판정한다. 만약 단계(S32)에서 패킷 손실율 γ이 50% 이하이면 단계(S33)로 진행하고, 패킷 손실율 γ이 50% 초과이면 단계(S34)로 진행한다.

[0101] 단계(S33)에서, 만약 연산된 복호화 성능 지수가 상위 포화 수준(sat_U) 이하이면 연산된 복호화 성능 지수의 값 대신에 상위 포화 수준(sat_U)을, 그렇지 않으면 연산된 복호화 성능 지수를 예측 복호화 성능 지수로서

출력한다.

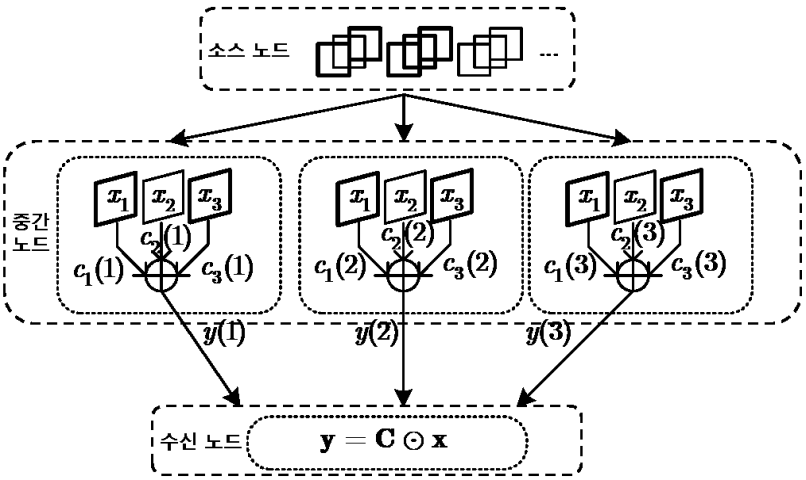
[0102] 단계(S34)에서, 만약 연산된 복호화 성능 지수가 하위 포화 수준(sat_L) 이하이면 연산된 복호화 성능 지수의 값 대신에 하위 포화 수준(sat_L)을, 그렇지 않으면 연산된 복호화 성능 지수를 예측 복호화 성능 지수로서 출력한다.

[0103] 이렇게 예측되는 복호화 성능을 기초로 근사 복호화를 위한 위치 유사성 정보 크기의 최적화를 시도할 수 있다.

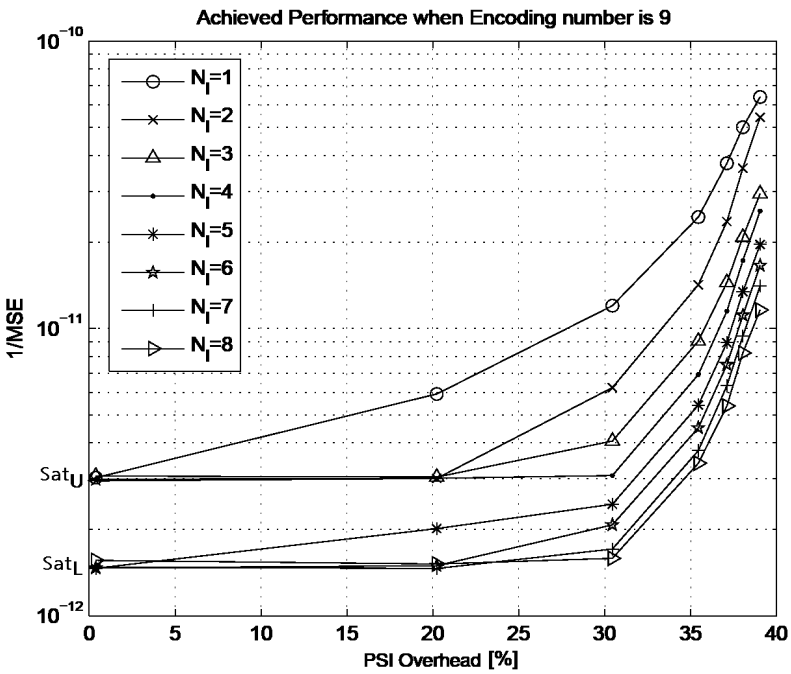
[0104] 이상과 같이 본 발명은 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명이 상기의 실시예에 한정되는 것은 아니며, 이는 본 발명이 속하는 분야에서 통상의 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 사상은 아래에 기재된 특허청구범위에 의해서만 파악되어야 하고, 이와 균등하거나 또는 등가적인 변형 모두는 본 발명 사상의 범주에 속한다 할 것이다.

도면

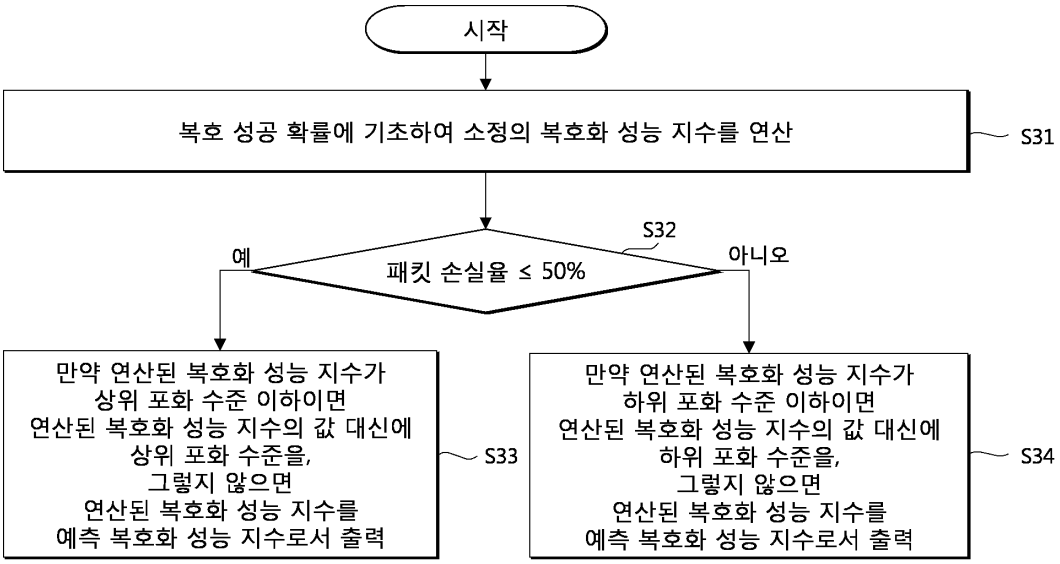
도면1



도면2



도면3


시작
복호 성공 확률에 기초하여 소정의 복호화 성능 지수를 연산
S31
S32
패킷 손실율 ≤ 50%
예
아니오
만약 연산된 복호화 성능 지수가
상위 포화 수준 이하이면
연산된 복호화 성능 지수의 값 대신에
상위 포화 수준을,
그렇지 않으면
연산된 복호화 성능 지수를
예측 복호화 성능 지수로서 출력
S33
만약 연산된 복호화 성능 지수가
하위 포화 수준 이하이면
연산된 복호화 성능 지수의 값 대신에
하위 포화 수준을,
그렇지 않으면
연산된 복호화 성능 지수를
예측 복호화 성능 지수로서 출력
S34