

인덱스 부호화에서 부호 길이 감소를 위한 보조정보 그래프 상의 필요충분조건

김재원, 곽희열, 김찬기, 노종선

서울대학교

{kjlw702, ghy1228, carisis}@ccl.snu.ac.kr, jsno@snu.ac.kr

A Necessary and Sufficient Condition of a Side-Information Graph for Index Coding to Reduce Codelength

Jaewon Kim, Kwak Heeyoul, Chanki Kim, and Jong-Seon No
Seoul National Univ.

요 약

본 논문에서는 보조정보를 이용하여 부호의 길이를 줄이는 인덱스 부호화 상황에서 부호의 길이를 줄이기 위하여 필요충분조건으로 필요한 보조정보 그래프의 특성에 대해 논한다. 또한 증명된 보조정보 그래프의 특성과 부호 길이와 관련하여 기존에 알려져 있던 조건을 비교한다.

I. 서 론

인덱스 부호화는 본래 위성통신을 위하여 도입된 개념이지만 다양한 정보이론적 문제들과의 연관성으로 인해 최근 학계에서 주목 받고 있다. 주어진 보조 정보 그래프를 이용하여 최적의 부호 길이를 찾아 내는 것이 인덱스 부호화의 주요 문제이며 최적의 부호 길이를 찾는 것은 NP-hard 문제로 알려져 있다 [1]. 이 외에 사이클(cycle)이 없는 보조정보 그래프에서의 최적의 길이는 어떠한 부호화도 시도하지 않을 때라는 것이 알려져 있다 [2].

일반적으로 사이클이 존재하지 않을 때 보조정보를 이용하여 부호의 길이를 줄일 수 없다는 것이 알려져 있지만 반대로 사이클이 존재하는 것이 항상 부호의 길이를 줄일 수 있는 것을 보장하지는 않는다. 즉 사이클이 존재하는 것은 인덱스 부호의 길이를 줄이기 위한 필요조건이다. 본 논문에서는 사이클과 유사한 보조정보 그래프의 특성을 찾고 이 그래프의 특성이 부호 길이를 줄일 수 있다는 사실과 필요충분조건임을 보인다. 그 다음으로 예시를 통하여 본 논문에서 제시한 보조정보 그래프의 특성과 기존에 알려져 있던 사이클의 성질을 비교한다. 이를 바탕으로 기존에 알려진 사이클 대신에 새로운 그래프 모양을 통하여 인덱스 부호화를 적용할 것인지를 판단할 수 있게 된다.

본 논문의 구성은 다음과 같다. II장에서는 인덱스 부호 및 본 논문의 표기법과 가정에 대해 간단히 소개한다. III장에서는 본 논문에서 제시하는 보조정보 그래프의 특성을 제시하며 주요 정리를 증명하고 마지막으로 IV장에서 결론을 맺는다.

II. 배경 이론

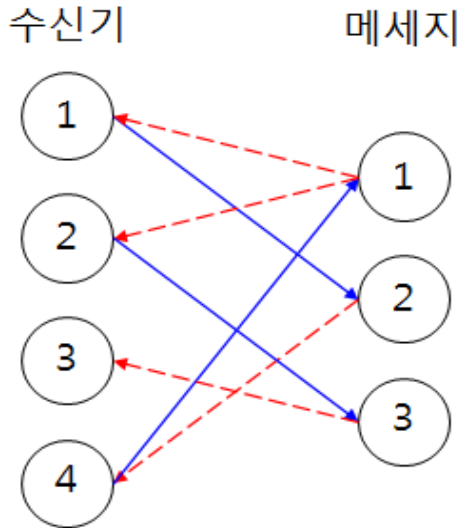
인덱스 부호화는 하나의 송신기와 다수의 수신기가 존재하는 브로드캐스트 채널(broadcast channel)을

가정한다. 인덱스 부호화가 다른 부호화 기법들과 다른 점은 송신기가 보조정보 그래프를 알고 있다는 것이다. 각 수신기들은 일부 정보들을 보조정보로 미리 알고 있으며 그 외의 다른 정보들의 일부를 송신기로부터 전달 받기를 원하는데 각 수신기들의 보조정보와 원하는 정보의 관계를 표현하는 것이 보조정보 그래프이다.

본 논문에서는 [3]의 표기법을 따른다. 메시지의 개수를 n 개, 수신기의 개수를 m 개로 생각한다. 송신기는 메시지 벡터 $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{F}_q^n$ 를 가지고 있으며 보조정보 그래프 G 를 알고 있다. 각 수신기는 이 중 하나의 정보를 수신 받기를 원하며 i 번째 수신기가 원하는 메시지의 인덱스를 $f(i)$ 로 정의하고 해당 원소를 $x_{f(i)}$ 로 정의한다. 또한, i 번째 수신기가 보조정보로 미리 알고 있는 메시지들의 인덱스집합을 X_i 로 정의한다.

일반적인 인덱스 부호화 상황에서 각 수신기가 원하는 메시지가 벡터이거나 여러 개의 메시지를 원하는 상황이 있을 수 있지만 해당 벡터의 사이즈가 정해져 있다면 위와 같은 가정으로 바꿀 수 있다는 것은 자명한 사실이다. 이와 더불어 본 논문에서는 채널에 에러가 없는 경우에 대해 다룬다.

<그림 1>은 메시지의 개수가 3 개이고, 수신기의 개수가 4 개인 보조정보 그래프의 예시이다. 보조정보 그래프는 수신기와 메시지의 2 개의 노드(node)로 이루어져 있으며 각 노드의 숫자는 수신기와 메시지의 인덱스를 의미한다. 메시지 노드로부터 수신기 노드로 가는 빨간 점선 화살표는 해당 인덱스의 수신기가 원하는 메시지의 인덱스를 나타낸다. 예를 들어, <그림 1>에서 첫 번째 메시지는 첫 번째와 두 번째 수신기가 수신 받기를 원하는 메시지이다. 수신기 노드에서 메시지 노드로 가는 파란 실선 화살표는 해당 인덱스의 수신기가 갖고 있는 보조정보의 인덱스를 의미한다. 예를 들어, <그림 1>에서 첫 번째 수신기는 두 번째 메시지를 보조정보로 미리 알고 있다.



<그림 1> $n = 3, m = 4$ 일 때 보조정보 그래프 예시

III. 본 론

본 논문의 결과를 증명하기에 앞서서, 인덱스 부호의 생성행렬을 A 라고 할 때 [3]번 논문의 Lemma 3.8 에 따르면 A 가 인덱스 부호의 생성행렬인 것은 아래의 식과 필요충분조건임이 알려져 있다.

$$wt(zA) \neq 0 \text{ for all } z \in \mathcal{L}(q, G) \quad (1)$$

이 때, $\mathcal{L}(q, G) = \{z \in \mathbb{F}_q^n : \exists i \in [m] \text{ s.t. } z_{x_i} = 0, z_{f(i)} \neq 0\}$ 로 정의되며, 이는 각 수신기에 대해 보조정보에 해당하는 성분은 0 이고 수신 받기를 원하는 성분은 0 이 아닌 벡터들을 모은 집합이다.

본 논문에서 제시하는 보조정보 그래프는 아래와 같은 집합 ψ 의 원소를 메시지의 집합으로 하며 원래 보조정보 그래프에서 이와 관련된 화살표와 수신기로 이루어진 부분그래프 G' 이다.

$$\psi = \{P \subseteq [n] \mid |X_i \cap P| \geq 1 \text{ for all } i \in [m] \text{ s.t. } f(i) \in P\} \quad (2)$$

Theorem 1: $\psi = \phi$ 와 인덱스 부호화로 부호 길이를 줄이지 못하는 것은 동치이다.

증명)

$\Rightarrow \psi = \phi$ 이면 메시지 인덱스들의 전체집합인 $[n]$ 도 원소가 아니므로 적어도 1 개의 수신기는 보조정보가 전혀 없다. 이러한 수신기의 인덱스를 i 라고 할 때 i 번째 수신기에 대해 $\mathcal{L}(q, G)$ 에 속하는 벡터들을 생각해 보면 보조정보가 전혀 없으므로 $f(i)$ 번째 성분만 0 이 아닌 모든 벡터들이 해당된다. 따라서 식 (1)을 생각하면 생성행렬 A 의 $f(i)$ 번째 행은 나머지 행들의 스패(span)에 속하지 않는다. 이로 인해 $x_{f(i)}$ 를 알고 있는 다른 수신기들에 대해 $x_{f(i)}$ 를 몰랐더라도 식 (1)을 생각하면 A 의 $f(i)$ 번째 행이 나머지 행들의 스패에 속하지 않으므로 결과가 같은 것을 확인할 수 있다. 따라서 보조정보 그래프에서 $x_{f(i)}$ 를 보조 정보로 미리 알고 있는 모든 수신기에 대해 $x_{f(i)}$ 로 가는 화살표를 모두 지우더라도 같은 문제를 푸는 것으로 생각할 수 있다. 이러한 화살표들을 모두 지우고 나면 $[n] \setminus \{f(i)\}$ 도 ψ 에 속하지

않으므로 수신 받기를 원하는 수신기 외에도 보조정보가 하나도 존재하지 않는 수신기가 적어도 1 개 존재한다. 위와 같은 논리로 이러한 수신기가 수신 받기를 원하는 인덱스에 해당하는 A 의 행도 나머지 행들의 스패에 속하지 않게 되고 결국엔 모든 A 의 행이 서로 선형적으로 독립이 되게 된다. 채널에 에러가 없는 경우이므로 인덱스 부호의 길이는 생성행렬 A 의 랭크(rank)와 같은 것이 자명하므로 $\psi = \phi$ 이면 부호의 길이는 인덱스 부호화로 부호 길이를 줄이지 못한 n 이 된다.

\Leftarrow 이 명제의 대우는 $\psi \neq \phi$ 이면 인덱스 부호화로 부호 길이를 줄일 수 있다는 것이다. $\psi \neq \phi$ 이면 임의의 원소 P 에 대해 $|P|-1$ 의 길이로 P 에 속하는 인덱스에 해당하는 원소를 원하는 수신기들이 각자가 원하는 것을 얻을 수 있다. 이유는 간단하게 P 에 속하는 인덱스에 해당하는 원소들을 $x_1, x_2, \dots, x_{|P|}$ 라고 한다면 $(x_1 + x_2, x_2 + x_3, \dots, x_{|P|-1} + x_{|P|})$ 를 인덱스 부호로 생각한다. 이 때, 각 수신기들이 $|P|$ 개의 메시지 중 적어도 1 개는 알고 있으므로 각자가 원하는 메시지를 얻어낼 수 있다.

예시로 <그림 1>을 보면 $\psi = \phi$ 임을 알 수 있다. 비록 보조정보 그래프에서 첫 번째 수신기로부터 네 번째 수신기를 거쳐 돌아오는 사이클이 존재하지만 $\psi = \phi$ 이므로 이 경우는 사이클이 있음에도 불구하고 인덱스 부호화로 부호 길이를 줄일 수 없게 된다.

구체적으로 설명하자면 <그림 1>의 보조정보 그래프에서 Theorem 1의 증명과정과 같이 수신기 노드에서 메시지 노드로 가는 화살표들을 지워나가면 2번, 4번, 1번째 수신기 순으로 모든 화살표가 지워지게 되며 이는 모든 수신기가 보조정보가 하나도 없는 경우와 같아지게 된다.

IV. 결 론

본 논문에서는 인덱스 부호화 상황에서 기존에 알려져 있던 사실인 보조정보 그래프에 사이클이 존재하지 않으면 인덱스 부호화로부터 부호 길이의 이득을 얻을 수 없다는 사실에서 나아가 인덱스 부호화로부터 부호 길이의 이득을 얻을 수 있는 것과 어떤 보조정보 그래프의 조건이 동치인지 증명하였다. 일반적으로 보조정보 그래프에서 사이클이 존재하더라도 부호 길이를 항상 줄이는 것을 보장하지는 못하지만 본 논문에서 제시한 보조정보 그래프의 조건을 만족하면 부호의 길이를 항상 줄일 수 있다는 것을 확인하였다. 이로부터 인덱스 부호화를 적용할지에 대한 새로운 기준을 제시하였다.

참 고 문 헌

[1] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1479-1494, 2011.

[2] M. J. Neely, A. S. Tehrani, and Z. Zhang, "Dynamic index coding for wireless broadcast networks," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7525-7540, 2013.

[3] S. H. Dau, V. Skachek, and Y. M. Chee, "Error correction for index coding with side information," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1517-1531, 2013.