JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지 Vol. 23, No. 10: 1195~1200, Oct. 2019

무잡음 그룹검사에 대한 확률적 검출 알고리즘

성진택*

A Probabilistic Detection Algorithm for Noiseless Group Testing

Jin-Taek Seong*

*Assistant Professor, Department of Convergence Software, Mokpo National University, Muan, 58554, Korea

요 약

본 논문은 그룹검사(Group Testing)에 대한 검출 알고리즘을 제안한다. 그룹검사는 다수의 샘플 중 극히 일부의 결함 샘플을 찾는 문제로써 이것은 압축센싱 문제와 유사하다. 본 논문에서는 잡음이 없는 그룹검사를 정의하고, 결함 샘플을 검출하기 위한 확률 기반의 알고리즘을 제안한다. 제안하는 알고리즘은 입력과 출력 신호 간 외부확률들이 서로 교환하여 출력 신호의 사후 확률이 최대가 되도록 구성한다. 그리고 검출 알고리즘에 대한 모의실험을 통해 그룹검사 문제에서 결함 샘플을 찾는다. 본 연구에 대한 모의시험 결과는 정보이론의 하한치와 비교하여 입력과 출력 신호 크기에 따라 실패확률이 얼마나 차이가 있는지 확인한다.

ABSTRACT

This paper proposes a detection algorithm for group testing. Group testing is a problem of finding a very small number of defect samples out of a large number of samples, which is similar to the problem of Compressed Sensing. In this paper, we define a noiseless group testing and propose a probabilistic algorithm for detection of defective samples. The proposed algorithm is constructed such that the extrinsic probabilities between the input and output signals exchange with each other so that the posterior probability of the output signal is maximized. Then, defective samples are found in the group testing problem through a simulation on the detection algorithm. The simulation results for this study are compared with the lower bound in the information theory to see how much difference in failure probability over the input and output signal sizes.

키워드 : 검출 알고리즘, 그룹검사, 실패확률, 최대 사후 확률

Keywords : Detection Algorithm, Group Testing, Failure Probability, Maximum a Posterior

Received 5 July 2019, **Revised** 15 July 2019, **Accepted** 29 July 2019

*** Corresponding Author** Jin-Taek Seong(E-mail:jtseong@mokpo.ac.kr, Tel:+82-61-450-2772)
Assistant Professor, Department of Convergence Software, Mokpo National University, Muan, 58554, Korea

Open Access **http://doi.org/10.6109/jkiice.2019.23.10.1195** print ISSN: 2234-4772 online ISSN: 2288-4165

Ⅰ. 서 론

그룹검사(Group Testing)는 Dorfman [1]에 의해 소개된 이후 반세기 동안 다양한 분야에서 그 활용 범위를 확장하였다. 2006년 정보이론(information theory) 분야에서 Donoho 교수의 압축센싱(Compressed Sensing) [2] 문제가 소개되면서 그룹검사는 다시 한 번 연구적인 의미와 가치를 되새기는 계기가 된다. 또한 압축센싱의 폭발적인 관심과 연구결과는 지난 10여년에 걸쳐 다양한 분야에서 수많은 문제를 해결할 수 있는 원동력으로 작용하게 된다. 최근 들어 그룹검사의 성능을 더 정교하게 살펴보고자 하는 움직임이 있었으며, 나아가서는 신호의 희소성(sparsity)에 따라 복원할 수 있는 성능 한계치를 밝히고 있다[3].

그룹검사는 2차 세계대전 중 미국 공중 보건국에서 모든 매독 남성을 찾기 위한 프로젝트에서 시작된다. 당시에 매독 검사는 군인 개개인의 혈액 샘플을 채취하여 매독의 감염 여부를 확인하였다. 그러나 매독 검사의 군인 수가 상당히 많았기 때문에 검사 비용이 막대하였고 뿐만 아니라 검사 시간 또한 많이 소요되어 새로운 검사 방법을 찾아야만 했다[4]. 이후 Dorfman에 의해 그룹검사가 처음으로 제안된다[1].

초기의 그룹검사는 다음과 같은 방법을 이용하여 매독 검사를 수행한다. 먼저 여러 군인들의 혈액 샘플을 한 곳에 섞어서 매독 성분에 반응하는지 확인한다. 그리고 그 결과가 양성(positive)일 때는 적어도 한 명 이상의 군인이 매독에 감염되었다는 것을 말해준다. 반대로 음성(negative)일 경우에는 매독 검사에 사용한 모든 혈액 샘플들이 모두 매독에 감염되지 않았음을 확인할 수 있다. 이와 같은 매독 검사가 가능한 이유는 대다수의 군인들은 매독에 감염되지 않고 오직 극소수의 일부 군인들만이 매독에 감염되었기 때문이다. 그래서 그룹검사 문제는 다음 두 가지에 대해 주로 다뤄지고 있다. 먼저 하나의 그룹(group)에 포함될 샘플들을 어떻게 생성해야 하는지에 대한 것이다. 두 번째는 다수의 샘플들 중에 결함이 있는 샘플을 어떠한 검출방법을 이용하여 찾아야 하는지에 관한 것이다.

본 논문에서는 그룹검사 문제를 다음과 같이 간단히 정의한다. N개의 샘플 중 D개의 샘플이 결함일 때, 이 결함 샘플을 찾기 위해 필요한 검사 수를 T라고 하자. 본 논문에서는 최대 사후 검출(maximum a posterior)을 이용하여 확률적인 검출 알고리즘을 제안한다. 그리고 샘플 크기 N과 D에 따른 성능 변화를 확인하고자 하며, 나아가서는 정보이론에서 얻은 성능 하한치(lower bound)와 비교를 한다.

논문 2장에서는 그룹검사와 관련된 연구 내용을 살펴보고, 3장에서는 본 연구에서 정의하고 있는 그룹검사 문제를 구체적으로 알아본다. 그리고 본 연구에서 제안하는 검출 알고리즘에 대한 내용을 본문 4장에서 설명하고 모의실험 결과를 살펴본다. 마지막으로 5장에서 본 연구로부터 얻은 실험결과에 대한 의미와 맺는말로 마무리 짓는다.

Ⅱ. 관련 연구

그룹검사에서 사용하는 검출 알고리즘은 Dorfman의 의해 처음으로 소개된 이후 수많은 알고리즘들이 제안되었다. 본 장에서는 그룹검사 문제에서 소개된 주요 검출 알고리즘들을 살펴본다.

먼저 알아 볼 검출 알고리즘은 이진 분할 알고리즘이다. 이 알고리즘은 기본적으로 최적의 적응형(adaptive) 그룹검사 알고리즘이라고 한다[4]. 이진 분할 알고리즘은 N개의 샘플에서 D개보다 작거나 같은 결함 샘플을 찾기 위한 알고리즘으로써 N과 D에 대한 크기에 따라 다음과 같은 동작과정으로 요약된다. 첫 번째 단계에서, $N \le 2D-2$인 경우, D개의 결함 샘플을 찾기 위해서 개별(individual) 검사를 수행한다. 그렇지 않으면, $L = N-D+1$과 $\alpha = \lfloor \log_2 L/D \rfloor$를 각각 설정한다. 두 번째 단계에서, 2^{α} 크기 만큼 그룹 단위로 묶어 그룹검사를 수행한다. 이때 검사의 결과가 음성이면, 그 그룹 내의 모든 샘플들은 모두 정상으로 판정하고 샘플 크기 $N = N-2^{\alpha}$로 재설정하여 첫 번째 단계를 수행한다. 그렇지 않으면, 이진 탐색을 수행하여 한 개의 결함 샘플과 정상 샘플 수 S에 대해 다음과 같이 재설정하고, $N = N-1-S$과 $D = D-1$, 첫 번째 단계를 수행한다.

일반적인 이진 분할 알고리즘은 $p > 0$에 대해 N과 D의 크기에 따라 다음과 같은 검사 수 T가 필요하다 [4].

$$T = \begin{cases} N & N \le 2D-2 \\ (\alpha+2)D+p-1 & N \ge 2D-1 \end{cases} \tag{1}$$

여기서 N/D이 클 경우, T는 $D\log_2(N/D)$에 값에 수렴한다[4]. 이것은 Li's S-stage 알고리즘에서 요구한 $T=(eD\log_2(N/D))/\log_2 e$과는 차이가 있다. 엄밀하게 말하면, 이진 분할 알고리즘은 다음과 같은 의미에서 최적값에 이르고 있다. $D \geq 2$일 때, $T-B_I(D,N) \leq (D-1)$을 충족하며, 여기서 $B_I(D,N)=\left\lceil \log_2 \sum_{i=0}^{D} \binom{N}{i} \right\rceil$를 말한다[4].

다음으로 COMP(Combinatorial Orthogonal Matching Pursuit) 알고리즘을 살펴보자[5]. 이 알고리즘은 단순한 비적응형(nonadaptive) 그룹검사 알고리즘이다. COMP 알고리즘은 다음과 같은 동작과정을 수행한다. 먼저 그룹행렬(group matrix)의 각 원소는 i.i.d.(independent and identically distributed) 확률 변수로써 $1/D$의 확률로 1을 갖고 $1-1/D$의 확률로 0을 갖는다. 검출방법은 열조합(column-wise)을 통해 진행된다. 그룹의 검사 결과가 양성이면, 그 샘플들은 결함이라고 판명한다. 반면에 검사 결과가 음성이면, 샘플들은 모두 정상이라고 판명한다. 같은 의미로, 하나의 샘플이 음성인 결과를 갖는 그룹에 포함되어 있으면, 그 샘플은 정상 샘플이다. 반면에 그 샘플이 양성인 결과를 갖는 그룹에 포함되어 있으면 그 샘플은 일단 결함 샘플이라고 가정한다.

COMP 알고리즘은 임의의 상수 $\epsilon > 0$에 대해 평균 에러 오율 $N^{-\epsilon}$보다 작거나 같도록 하기 위해 필요한 검사 수 T는 다음과 같다[5].

$$T \geq eD(1+\epsilon)\ln(N) \tag{2}$$

여기서 검사 수는 평균 에러 오율에 대해 하한치의 상수배 e를 말한다.

DD(Definite Defectives) 알고리즘은 어떠한 false positive 없애고자 하는 노력의 일환으로 COMP 알고리즘을 확장한 형태이다[6]. DD 알고리즘의 성능은 COMP 알고리즘의 성능을 개선하였다. DD 알고리즘의 검출방법은 COMP 알고리즘의 유용한 속성을 이용한다. COMP 알고리즘에서 정상이라고 판단한 샘플은 false negative 없이 확실하게 정상 판정을 한다. 그리고 다음과 같은 DD 알고리즘의 동작과정을 수행한다.

첫 번째 단계에서는 COMP 알고리즘과 동일하게 동작한다. 이후 정상 샘플들은 모두 제거하고 남아 있는 샘플들은 결함이나 정상 상태 "possibly defective"를 지닌다. 두 번째 단계에서는 하나의 샘플만 참여하여 양성 결과가 나왔다면, 그 샘플은 결함이 있는 상태이다. 마지막 단계에서는 나머지 다른 샘플들은 모두 정상 상태라고 판명한다. 이러한 이유는 결함 샘플 수는 전체 샘플 수에 비해서 매우 작다는 점에서 찾을 수 있다. 여기서 첫 번째와 두 번째 단계는 정확히 정상 샘플과 결함 샘플을 판단한다. 반면에 마지막 단계에서는 결함일 수도 있는데도 불구하고 모두 정상 샘플이라고 판단함에 따라 오류가 발생한다. 그러므로 DD 알고리즘은 단지 false negatives만을 발생시킨다.

SCOMP(Sequential COMP)은 DD 알고리즘이 마지막 단계에서까지 오류를 일으키지 않는다는 이점을 이용한 알고리즘이다[6]. 여기서 남아 있는 모든 샘플들은 모두 정상이라고 가정한다. 결함으로 판단된 샘플의 집합을 K라고 하자. 만약 검사가 집합 K에서 하나의 샘플이라도 포함하고 있으면 양성 결과를 나타낸다. DD 알고리즘에 의해 판단된 결함 샘플의 집합은 모두 양성 결과를 일으킨다고 말할 수 없다. 이것은 명확히 판명할 수 없는 검사 결과는 반드시 하나의 은닉 결함 샘플을 포함하고 있어야 한다. SCOMP 알고리즘을 이용한 모의실험을 통하여 최적값에 근접한 결과를 보여 주었다[6].

Ⅲ. 그룹검사 문제

본 장에서는 그룹검사 문제에 대해 구체적인 정의한다. 먼저 입력 신호 $X=(x_1,x_2,\cdots,x_N)$는 크기가 N인 이진 벡터이다, 여기서 $X\in\{0,1\}^N$. 여기서 j번째 샘플이 결함일 경우, $x_j=1$로 표현되며, 그렇지 않을 경우 $x_j=0$으로 정의된다. 그러므로 입력 X의 각 원소는 0 또는 1로 표현된다. 본 논문에서는 벡터 $X=(x_1,x_2,\cdots,x_N)$의 각 원소가 다음과 같은 (i.i.d.) 베르누이 확률분포를 갖는 것으로 가정한다,

$$\Pr\{x_j=\theta\}=\begin{cases}1-\delta & \text{if } \theta=0\\ \delta & \text{if } \theta=1\end{cases} \tag{3}$$

여기서 $\delta=D/N$는 감염률을 말하며 θ는 임의의 변수이다.

그룹행렬 $A\in\{0,1\}^{T\times N}$는 T개의 행과 N개의 열을 갖는다. 이때 $i\in\{1,2,3,\ldots,T\}$ 번째 그룹(group)에 $j\in\{1,2,3,\ldots,N\}$ 번째 샘플을 포함하면 이에 대응하는

그룹행렬의 원소 $A_{ij}=1$로 표현되며, 그렇지 않으면 $A_{ij}=0$으로 표현된다. 다시 말하면, 각 그룹행렬의 원소가 1일 경우에는 해당 열을 의미하는 j번째 샘플을 포함하여 그룹검사를 수행한다[4].

다음은 그룹검사의 수학적인 표현을 좀 더 자세히 설명한다. 앞서 정의한 벡터 X와 그룹행렬 A는 다음과 같은 그룹검사로 정의된다,

$$Y=A\oplus X \tag{4}$$

여기서 Y는 검사결과 벡터이며, i번째 그룹검사에서 양성이면 $y_i=1$이고, 그렇지 않으면 0이다. 그리고 $\oplus$은 **AND**와 **OR**로 처리되는 논리연산을 말한다. 다음 식 (5)는 간단한 그룹검사 $Y=A\oplus X$를 보여준다.

$$\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}0\,1\,1\\0\,1\,1\\1\,0\,0\end{bmatrix}\oplus\begin{bmatrix}1\\0\\0\end{bmatrix} \tag{5}$$

식 (5)에서 Y의 첫 번째 값은 (0 1 1)$\oplus$(1 0 0) = (0 **AND** 1) **OR** (1 **AND** 0) **OR** (1 **AND** 0) = 0 음성 결과를 얻는다. 같은 방법으로 검사결과 벡터 Y의 모든 값을 얻는다. 예제에서 보여준 바와 같이 각 그룹검사에 포함된 벡터 X의 원소가 한 개라도 1이면 그 검사결과 값은 1이다. 역으로 그룹검사를 수행하는 벡터 X의 모든 원소들이 0이면 그 결과는 음성이다.

Ⅳ. 결함 샘플 검출

4.1. 검출 알고리즘

본 장은 그룹검사 문제에 대한 확률적 검출 알고리즘을 제안한다. 검출 알고리즘은 최대 사후 검출방법을 이용하여 설계한다. 먼저 최대 사후 검출에 대한 이해를 시작으로 검출 알고리즘을 전개한다. 신호 벡터 X는 앞서 정의한 바와 같이 식 (3)을 따른다. 우리는 주어진 검사결과 Y와 그룹행렬 A를 이용하여 신호 X가 최대 확률을 갖게끔 검출한다. 이를 수학적으로 표현하면 다음과 같이 정의된다,

$$\begin{aligned}\hat{X}&=\arg\max \Pr\{X|Y\}\\&=\arg\max\prod_{j=1}^{N}\Pr\{x_j|Y\}\end{aligned} \tag{6}$$

여기서 신호 X의 각 원소가 각각 독립이라고 가정한다. 그러므로 식 (6)은 다음과 같이 원소 x_j의 사후 확률로 표현된다,

$$\begin{aligned}\Pr\{x_j|Y\}&=\sum_{X\setminus\{x_j\}}\Pr\{X,Y\}\\&=\sum_{X\setminus\{x_j\}}\Pr\{Y|X\}\Pr\{X\}\\&=\sum_{X\setminus\{x_j\}}\prod_{i=1}^{T}\Pr\{y_i|X\}\prod_{j=1}^{N}\Pr\{x_j\}\end{aligned} \tag{7}$$

여기서 기호 $\setminus$은 여집합을 말하고, 각 그룹검사가 모두 독립일 경우 마지막 등식은 만족한다.

다음으로 본 논문에서 제안하는 검출 알고리즘의 동작원리를 살펴본다. 먼저 동작원리를 설명하기에 앞서서 식 (4)에 대해 그래프 표현방법을 다음과 같이 소개한다. 그림 1은 신호 X와 Y, 그리고 그룹행렬 A를 이용하여 그래프로 표현한다.

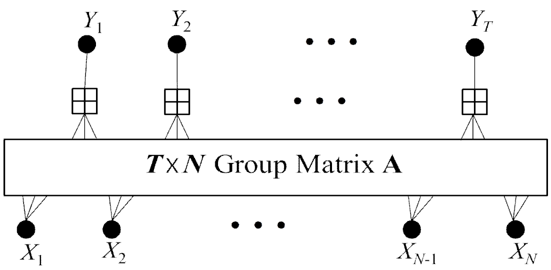

Fig. 1 Graph representation with edges and nodes for the group matrix A, the input and output signals X and Y

그래프에서 각각의 엣지(edge)는 i번째 그룹검사에서 j번째 샘플을 포함할 경우 해당 그룹행렬의 원소 $A_{ij}=1$인 경우에 형성된다. 그렇지 않으면 엣지는 존재하지 않는다. 다시 말하면, 임의의 신호 y_i와 x_j 사이에 엣지가 형성된 경우 그룹행렬의 원소 A_{ij}가 1이다. 또한 그룹행렬의 각 행과 각 열에 동일한 1의 수를 갖는 경우 상수 행(constant row)과 상수 열(constant column) 그룹행렬이라 부른다.

다음으로 결함 샘플 검출을 위한 알고리즘을 설명하고자 한다. 본 논문에서 제안하는 알고리즘은 기본적으로 채널코딩(channel coding) 이론에서 발표한 LDPC 코드의 sum-product 알고리즘을 기반으로 설명된다[7]. 이 알고리즘은 정보이론에서의 한계치에 근접한 성능을 보여주었다. Mackey 방법과 제안하는 알고리즘과의 차

이점은 연산과정에서 드러난다. 채널코딩의 경우 유한체(Finite Fields)에서 연산을 수행하지만, 그룹검사는 논리연산을 기반으로 수행한다는 측면에서 다르다. 그러므로 채널코딩에서는 패러티-체크 과정이 콘볼루션 연산을 통한 과정인 반면 그룹검사에서는 그렇지 않다는 것이다. 그리고 그룹검사에서 잡음(noise)이 있는 경우에 대한 연구는 [8]-[10]에서 다루고 있다. 본 연구에서는 잡음이 없는 것으로 가정한다.

본 논문에서 제안하는 검출 알고리즘의 동작은 크게 두 개의 외부확률(extrinsic probability)이 반복적으로 정보를 주고받는 과정으로 설명된다. 다시 말하면 식 (7)에서 신호 x_j의 사후 확률이 최대가 되도록 노드 y_i와 x_j 사이에 형성된 엣지에 대한 외부확률의 반복적인 교환을 말한다. 각 외부확률은 각각 노드 x_j에서 y_i로 전달하는 확률 ξ_{ji}과 노드 y_i에서 x_j로 전달하는 확률 δ_{ij}로 정의한다. 즉 그룹검사 그래프의 엣지가 존재하는 경우, 그룹행렬의 원소가 0이 아닐 때, 두 개의 외부확률은 서로 교환한다.

다음은 신호 X의 결함 샘플을 검출하기 위한 자세한 알고리즘을 살펴본다. 검출 알고리즘은 반복 과정을 통해 식 (7)의 확률이 최대가 되도록 설계된다. 크게 초기화 단계, 외부확률 ξ_{ji}과 δ_{ij} 산출, 식 (4)의 조건을 충족하는지 $Y = A \oplus X$ 검증 과정을 걸쳐 알고리즘이 동작한다. 초기화 단계에서는 그룹행렬 A와 식 (3)에 의한 신호 x_j에 따른 결과값 y_i를 설정한다. 본 논문에서는 상수 열 그룹행렬을 이용한다. 그리고 신호 x_j의 식 (3)의 확률분포를 ψ_x라고 정의한다. 검출 알고리즘의 첫 단계에서는 노드 x_j는 확률분포 ψ_x을 이용하여 외부확률 ξ_{ji}을 산출하여 노드 y_i에 전달한다. 여기서 초기 외부확률 δ_{ij}은 0과 1에 대한 확률이 같다고 가정한다. 다음 단계에서 노드 y_i는 x_j로부터 전달된 외부확률 ξ_{ji}을 이용하여 외부확률 δ_{ij}를 다음과 같이 구한다,

$$\delta_{ij}(\theta) = \sum_{\{X: y_i = A_i \oplus X\}} \prod_{j' \in T(i) \setminus \{j\}} \xi_{j'i}(\theta) \tag{8}$$

여기서 $T(i) = \{j : A_{ij} \neq 0\}$은 그룹행렬의 i 행에 대해 0이 아닌 값에 대응하는 j 열의 집합이고, 이때 자신의 엣지로부터 생성된 외부확률 ξ_{ji}를 제외하고 계산된다. 또한 다음의 $y_i = A_i \oplus X$ 조건을 만족하는 경우에만 식 (8)을 적용한다.

외부확률 ξ_{ji}는 다음과 같이 구한다,

$$\xi_{ji} = \lambda \psi_x(\theta) \prod_{i' \in \Im(j) \setminus \{i\}} \delta_{i'j}(\theta) \tag{9}$$

여기서 $\Im(j) = \{i : A_{ij} \neq 0\}$은 그룹행렬의 j 열에 대해 0이 아닌 값에 대응하는 i 행의 집합이고, 변수 λ를 사용하여 확률분포를 정규화한다. 결과적으로 신호 x_j의 최대 사후 확률을 $\xi_j(\theta)$으로 정의하고, 식 (7)에서 유도한 바와 같이 외부확률 $\xi_{ji}(\theta)$을 이용하여 신호 x_j의 최대 사후 확률은 다음과 같다,

$$\xi_j(\theta) = \psi_x(\theta) \prod_{i \in \Im(j)} \xi_{ji}(\theta) \tag{10}$$

검출 알고리즘은 식 (8)과 (9)을 반복적으로 계산하여 외부확률을 구한 후, 식 (4) $Y = A \oplus X$ 조건을 만족하는지 검증한다. 이때 조건이 충족하면 알고리즘의 반복 수행을 멈추고, 그렇지 않은 경우 설정된 반복 횟수만큼 계속해서 수행한다.

4.2. 모의실험

본 장에서는 그룹검사 문제에 대해 확률적 검출 알고리즘의 성능을 살펴본다. 그리고 성능평가 결과를 그림 2에서 보여준다. 본 논문의 검출 알고리즘으로 사용된 모의시험은 최대 50번의 반복 횟수를 사용한다. 그리고 논문 [7]에서 소개한 패러티-체크 행렬을 그룹행렬로써 사용한다. 제안한 검출 알고리즘의 성능을 확인하기 위해 신호 X의 길이는 1000, $N = 1000$으로 설정하여 모의시험을 진행하였다. 본 모의시험에서는 그룹행렬의 상수 열(constant column)이 5일 때 그룹검사의 검출 성능에 대한 실패확률(failure probability)을 얻는다.

그림 2는 총 1000개의 샘플 중 결함 샘플이 50개인 경우($D = 50$) 모의시험 한 결과를 보여준다. 그림 2에서 보여주는 바와 같이 그룹검사의 검사 수 T에 따라 실패확률을 얻는다. 특히 검사 수를 많이 할수록 실패확률은 낮아지지만, 반대로 작을 경우 실패확률 높아짐을 확인할 수 있다. 또한 논문 [6]에서 도출한 정보이론의 성능 하한치는 $N = 1000$이고 $D = 50$인 경우 $T = 285$를 얻는다. 본 논문에서 제안하는 검출 알고리즘의 성능 결과와 성능 하한치와 비교했을 때 그 차이가 크지 않음을 확인할 수 있다. 그림 2를 통해 우리는 그룹검사에서 일정한 결함을 갖는 신호를 검출하기 위해 필요한 검사 수가 얼마만큼 커야 하는지에 대한 답을 말해주고 있다.

그림 2에서 알 수 있듯이 검출의 실패확률 0.01 이하 달성에 필요한 그룹검사의 검사 수 T는 380이상에서 대부분의 결함 샘플을 찾는다.

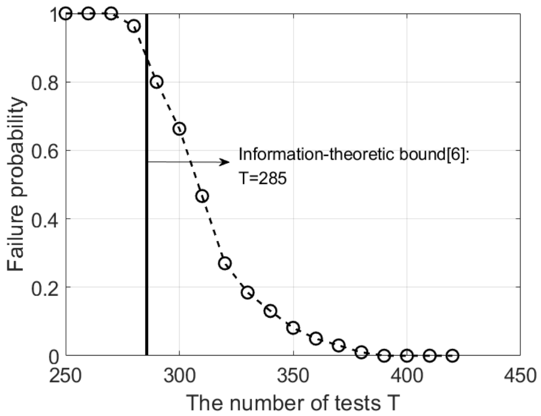

Fig. 2 Simulation results of the failure probability with $N=1000$ and $D=50$ compared with [6]

Ⅴ. 결 론

본 논문은 잡음이 없는 그룹검사를 정의하였으며, 또한 결함 샘플을 검출하기 위한 확률 기반의 알고리즘을 제안했다. 검출 알고리즘은 기본적으로 최대 사후 확률 방법을 이용하고 있으며, 입력과 출력 신호 간에 외부확률들이 서로 교환하여 출력 신호의 사후 확률이 최대가 되도록 구성했다. 모의실험 결과, 정보이론의 성능 하한치와 큰 차이가 없으며, 제안하는 검출 알고리즘을 통해 그룹검사 문제에서 결함 샘플을 찾을 수 있음을 확인했다.

ACKNOWLEDGEMENT

This Research was supported by Research Funds of Mokpo National University in 2018.

References

[1] D. Robert, “The Detection of Defective Members of Large Populations,” *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 436-440, Dec. 1943.

[2] D. L. Donoho, “Compressed Sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

[3] V. Ganditota, E. Grigorescu. S. Jaggi, and S. Zhou, “Nearly Optimal Sparse Group Testing,” *IEEE Transactions on Information Theory*, vol. 65, no. 5, pp. 2760-2773, May. 2019.

[4] D. Z. Du, and F. K. Hwang, *Pooling Designs and Nonadaptive Group Testing: Important Tools for DNA Sequencing*, World Scientific, 2006.

[5] C. L. Chan, P. H. Che, S. Jaggi, and V. Saligrama, “Non-adaptive probabilistic group testing with noisy measurements: near-optimal bounds with efficient algorithms,” *49th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1832 - 1839, Sep. 2011.

[6] M. Aldridge, L. Baldassini, and O. Johnson, “Group Testing Algorithms: Bounds and Simulations,” *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3671-3687, Jun. 2014.

[7] M. C. Davey, and D. Mackey, “Low-density parity-check codes over GF(q),” *IEEE Communications Letters*, vol. 2, no. 6, pp. 165-167, Jun. 1998.

[8] M. Aldridge, “The Capacity of Bernoulli Nonadaptive Group Testing,” *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7142-7148, Nov. 2017.

[9] J. Scarlett, and V. Cevher, “Near-Optimal Noisy Group Testing via Separate Decoding of Items,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 5, pp. 902-915, Oct. 2018.

[10] J. Scarlett, “Noisy Adaptive Group Testing: Bounds and Algorithms,” *IEEE Transactions on Information Theory*, vol. 65, no. 6, pp. 3646-3661, Jun. 2019.

성진택(Jin-Taek Seong)

2014년 광주과학기술원 정보통신공학과 공학박사
2008년~2010년 LG전자 주임연구원
2014년~2016년 대구경북첨단의료산업진흥재단 연구원
2016년~2017년 방위사업청 방송통신주사
2017년~2018년 호남대학교 정보통신공학과 조교수
2018년~현재 목포대학교 융합소프트웨어학과 조교수
※관심분야 : 정보이론, 인공지능, 코딩이론