

효율적인 코로나19 진단을 위한 그룹검사 체계

성진택*

Group Testing Scheme for Effective Diagnosis of COVID-19

Jin-Taek Seong*

요약 최근 코로나19 확산과 피해가 늘어나는 가운데 감염을 차단하기 위해 가장 중요한 것은 감염자를 조기에 찾아내는 것이다. 지난 반세기 전에 등장한 그룹검사(group testing)가 최근 코로나19 진단 방법으로써 활용 가능하며 매우 효율적인 방법으로 자리 잡고 있다. 본 논문에서는 기존의 그룹검사 알고리즘들의 동작원리를 살펴본다. 그리고 압축센싱(compressed sensing)에서 제안한 희소 신호 복원 방법을 개선하여 그룹검사의 해법으로 제시한다. 압축센싱과 그룹검사는 연산 방법에서 차이가 있지만 희소 신호를 찾는다는 점에서 유사하다. 시뮬레이션 결과를 통해 제안하는 희소 신호 복원 방법의 성능 우수성을 보여준다. 주목할 점은 모든 결합 샘플을 정확히 찾고자 하는 그룹검사 시스템에서는 제안하는 방법이 다른 알고리즘보다 월등한 성능 향상을 보여준 것이다. 또한 결합 샘플 수가 적을 때보다는 많을 때 그 성능이 크게 개선된다.

Abstract Due to the recent spread and increasing damage of COVID-19, the most important measure to prevent infection is to find infected people early. Group testing which introduced half a century ago, can be used as a diagnostic method for COVID-19 and has become very efficient method. In this paper, we review the fundamental principles of existing group testing algorithms. In addition, the sparse signal reconstruction approach proposed by compressed sensing is improved and presented as a solution to group testing. Compressed sensing and group testing differ in computational methods, but are similar in that they find sparse signals. The our simulation results show the superiority of the proposed sparse signal reconstruction method. It is noteworthy that the proposed method shows performance improvement over other algorithms in the group testing schemes. It also shows performance improvement when finding a large number of defective samples.

Key Words : Group Testing, Compressed Sensing, Sparse Recovery, Diagnosis of COVID-19

1. 서론

코로나19로 인한 대규모 전염병 확산을 억제하기 위한 최선의 방법 중의 하나는 코로나 감염자를 빨리 찾아내는 것이다. 이를 위해 개별 코로나 검사를 실시하기 보다는 대규모 검사를 통해 감염자를 찾는 방법으로써 1943년 Dorfman에 의해 소개된 그룹검사(Group Testing)[1] 방법이 최근에 관련 학계에서 주목받고 있다[2]-[4]. 그룹검사는 개별 표본 집합을 하나의 그룹으로 묶어서 검사하는 방법으로써 각 개별 샘플검사가 아닌 그룹 또는 풀을 검사한다. 코로나19처

럼 발병률이 낮은 경우(실제 코로나19 발병률은 2% 내외 수준임 [17])에 몇 개의 샘플들은 묶어 하나의 그룹으로 검사할 수 있게 됨에 따라 진단에 필요한 총 검사 수를 상당히 줄일 수 있다[5].

그룹검사 문제를 보다 간단히 말하면 다음과 같다. N 개의 샘플 중 D 개의 샘플이 희소한 결함(sparse defect)을 갖는다고 가정한다. 이때, 모든 결함 샘플을 정확히 찾기 위해 필요한 검사 수 T 가 얼마 이상이 되어야 하는지를 알고자 하는 것이다. 개별검사 방법이라면 모든 샘플 N 을 하나씩 검사하기 때문에 검사 수

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (NRF-2020R111A307173912).

*Department of Convergence Software, Mokpo National University (jtseong@mokpo.ac.kr)

Received December 07, 2021

Revised December 08, 2021

Accepted December 17, 2021

T 는 N 과 같다, $T=N$. 그러나 그룹검사 방법을 이용하면 검사 수 T 는 N 보다 훨씬 작다 점에서 경제적이다, $T \ll N$.

본 논문에서는 그룹검사 문제에 대해 압축센싱(compressed sensing)[6]에서 제안한 희소 신호 복원 방법을 이용하여 결합 샘플을 찾는 알고리즘을 제안한다. 압축센싱에서는 실수연산(real-valued operation)으로 선형 역문제(inverse problem)를 다루고 있지만, 그룹검사에서는 논리연산(logical operation)을 통해 희소 신호를 찾는 문제를 고려한다. 또한 기존의 그룹검사 문제를 해결하기 위해 발표한 COMP(Combinatorial Orthogonal Matching Pursuit)[7], DD(Definite Defectives)[8]과 SCOMP(Sequential COMP)[8] 알고리즘들을 살펴보고, 제안하는 희소 신호 복원 방법과의 성능을 비교하여 그 결과를 평가한다. 제안하는 희소 신호 복원 방법은 기본적으로 압축센싱에서 소개된 L1-norm 최소화 기법을 따르며 최적화 문제를 전개하기 위해 일부 변형하여 사용한다.

본 논문은 다음과 같이 구성된다. 2장에서는 그룹검사 문제를 정의하고 그룹검사 문제에 대한 이론적인 경계값을 살펴보고, 또한 기존의 발표된 그룹검사 알고리즘들을 설명한다. 3장에서는 그룹검사 문제의 그룹행렬 생성 방법을 설명하며 희소 신호 복원 방법을 구체적으로 제안한다. 그리고 컴퓨터 시뮬레이션을 통해 기존의 알고리즘들과 제안하는 알고리즘과의 성능을 비교 평가한다. 마지막 결론에서 그룹검사 성능평가에 대한 결론과 맺는말로 마무리 짓고자 한다.

2. 그룹검사 정의 및 성능

2.1 그룹검사 문제

다음은 그룹검사 문제를 정의한다. 먼저 입력 신호 $X=(x_1, x_2, \dots, x_N)$ 는 크기가 N 인 이진 벡터이다, $X \in \{0,1\}^N$, $j \in \{1, 2, \dots, N\}$. 여기서, j 번째 샘플이 결합이 있다면, $x_j=1$, 그렇지 않으면 $x_j=0$. 본 논문에서는 X 의 0이 아닌 원소의 수를 K 라고 가정한다, $\|X\|_0 = K$. 그룹행렬 $A \in \{0,1\}^{T \times N}$ T 개의 행과 N 개의 열을 갖는다. 그룹행렬 A 는 i 번째 검사

A_i 에서 j 번째 샘플 x_j 을 포함하여 그룹검사를 실시하면 $A_{ij}=1$ 로 표현되고 그렇지 않으면 $A_{ij}=0$. 즉, 그룹행렬은 i 번째 검사 A_i 에서 j 번째 샘플 x_j 을 포함하여 그룹검사를 수행하는지 여부에 따라 행렬의 각 원소값이 1(참여) 또는 0(미참여)로 대응된다[5].

그룹검사 문제는 다음과 같이 벡터 X 와 그룹행렬 A 를 이용하여 수식으로 표현된다.

$$Y=AX \quad (1)$$

여기서 Y 는 결과 벡터이며, AX 의 행렬 연산은 일반적인 행렬 곱셈을 의미하는 것이 아니라 논리연산을 사용한다. 다음 예제와 같이 결과 벡터 Y 는 그룹행렬과 입력 신호의 행렬 연산, AX 의 행렬 연산은 다음과 같이 수행한다.

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

첫 번째 원소 y_1 은 $(1 \ 0 \ 0) \cdot (1 \ 0 \ 0)^T = (1 \text{ AND } 1) \text{ OR } (0 \text{ AND } 0) \text{ OR } (0 \text{ AND } 0) = 1$, $y_1=1$ (결과가 양성). 다른 원소도 같은 방법을 이용하면 $y_2=0$ (결과가 음성). 식(2)에서 보여준 바와 같이 그룹검사 결과는 검사에 참여하는 원소값 1이 하나라도 존재하면 그 결과값은 1로 표현된다. 그렇지 않고 모든 검사에 참여하는 원소값이 모두 0일 때는 그 결과값은 0이 된다. 식(2)의 경우, 개별검사를 수행하게 되면 3번을 검사해야 되지만, 2번의 그룹검사를 수행하여 정확히 1개의 결합 샘플을 찾게 되어 검사 수를 줄일 수 있다. 결국 그룹행렬 설계와 복원 방법에 따라 그룹검사 수 T 에 영향을 주게 된다.

2.2 그룹검사 경계값

다음은 그룹검사 문제에 대해 검사 수 T 가 얼마나 커야지 원래의 신호 X 를 찾을 수 있는지에 대한 이론적인 경계값을 보여준다. 어려움이 0에 근접하는 그룹검사 알고리즘에 대해 N 개의 샘플에서 D 개의 결합 샘플을 찾기 위해 필요한 검사 수는 정보 이론을 이용하여 그 경계값을 찾는다[5]. 이론적인 경계 하한치(lower bound)의 검사 수 T 는

$$T \geq \log_2 |S|, \quad (3)$$

여기서 S 는 샘플 공간이고 $|S|$ 은 집합 S 의 크기를 말한다. 검사 수 T 에 대해 D 개의 결합 샘플을 찾을 때, 성공 확률 P_s 에 대한 경계 상한치(upper bound)를 얻는다[9]. 성공 확률 P_s 은 다음의 상한치보다 작다.

$$P_s \leq T \log_2 \left(\frac{N}{D} \right)^{-1} \quad (4)$$

최근 우리는 정보이론에서 이용한 Fano's inequality 정리[10]를 이용하여 그룹검사의 에러율의 경계 하한치를 얻었다[11].

정리 1 [11]: 임의의 그룹행렬과 디코더에 대해 신호 희소율 $\gamma(=D/N)$ 이 $0 < \gamma \leq 0.5$ 일 때, 에러율 P_e 은 다음과 같다,

$$P_e \geq H_b(\gamma) - \frac{T}{N} - \frac{1}{N} \quad (5)$$

여기서 $H_b(\gamma)$ 는 γ 값을 갖는 이진 엔트로피를 말한다.

2.3 COMP 알고리즘

COMP 알고리즘은 다른 복잡한 알고리즘의 기본이 되는 단순한 비적응형 그룹검사 알고리즘이다[7]. COMP 알고리즘은 다음과 같은 동작과정으로 설명된다. 먼저 그룹행렬의 각 원소는 $1/D$ 의 확률로 1을 갖고 $1-1/D$ 의 확률로 0이다. 디코딩은 열 조합(column combination)을 통해 수행된다. 그룹행렬의 하나의 열과 결과 벡터 값을 비교를 통해 해당 열에 대응하는 x_j 를 찾는 방법이다. 먼저 그룹행렬의 열 내에 1이 있는 위치의 대응하는 행에 대한 결과값이 1이면, 그 결과 벡터는 1이어야 한다. 그렇지 않으면 결과 벡터값을 0으로 판단한다. 결국 COMP 알고리즘은 위음성(false negative)을 발생하지 않으며, 반면에 결합 샘플이라고 판정한 것 중에는 정상 샘플이 포함되어 있기 때문에 위양성(false positive)이 발생된다. COMP 알고리즘은 평균 에러 오율이 $N^{-\delta}$ 보다 작거나 같도록 하기 위해 필요한 검사 수 T 는 다음을 만족한다[7],

$$T \geq eD(1+\delta)\ln(N) \quad (6)$$

여기서 검사 수 T 는 평균 에러 오율에 대해 하한치의 상수배를 말한다.

2.4 DD 알고리즘

그룹검사에서는 위양성과 위음성을 최소화할 필요가 필요하며 이를 해결하고자 DD 알고리즘이 제안된다[8]. DD 알고리즘은 위양성을 제거하기 때문에 COMP 알고리즘 보다 성능이 우수하다[8]. DD 알고리즘의 디코딩은 COMP 알고리즘의 유용한 속성을 이용한다. COMP 알고리즘을 통해 정상 샘플은 위음성 없이 정확히 찾을 수 있다. 먼저 COMP 알고리즘을 수행하여 정상 샘플들은 모두 찾아낸다. 남아 있는 샘플들은 결합 또는 정상 중 하나에 해당하는 결합 가능한 상태라고 할 수 있다. 이어서 하나의 샘플만 그룹검사에 참여하여 결과가 양성이면 그 해당 샘플은 결합이 있는 것으로 판단한다. 같은 방법을 통해 하나의 샘플만 그룹검사에 참여한 모든 샘플의 상태를 찾는다. 이후 정상인지 결합인지 판단할 수 없는 샘플들은 모두 정상 상태라고 예측한다. 이렇게 판단할 수 있는 이유는 결합 샘플 수는 전체 샘플 수에 비해 매우 작기 때문에 가능하다. 물론 정상 샘플이라 판단한 것들 중에는 결합이 있을 수 있기 때문에 DD 알고리즘은 위음성만 발생된다.

2.5 SCOMP 알고리즘

DD 알고리즘이 위양성을 발생시키지 않는 장점을 확장하여 SCOMP 알고리즘이 개발되었다[8]. DD 알고리즘에서는 최종적으로 남아 있는 모든 샘플들은 모두 정상이라고 가정하였다. 결합으로 판단된 샘플의 집합을 Ω 라고 가정하자. 만약 그룹검사가 집합 Ω 에서 하나의 샘플이라도 포함하고 있으면 양성 결과를 나타낸다. DD 알고리즘에서 판단한 결합 샘플 집합은 모두 양성이라고 확신할 수 없다. 결국 확실히 판단할 수 없는 검사 결과는 반드시 하나 이상의 은닉 결합 샘플을 포함함을 의미한다.

SCOMP 알고리즘은 다음과 같은 단계로 동작한다. 먼저 DD 알고리즘을 수행하여 집합 Ω 를 찾아 초기 결합 샘플 집합으로 정의한다. 그리고 집합 Ω 에 대해

각 샘플에 대해 양성 결과를 추적할 수 있으면 종료한다. 이후 집합 Ω 을 갱신된 결합 샘플 집합으로 정의한다. 만약 집합 Ω 가 정확히 판단할 수 없는 검사 결과에 대해 불명확한 결합 가능한 샘플로 결정한다. 마지막으로 이것을 결합 샘플로 판단한다. 위의 과정을 반복 수행하여 디코딩 과정을 종료한다. SCOMP 알고리즘은 시뮬레이션 결과를 통하여 최적화에 근접한 결과를 보여 주었다[8].

3. 그룹검사의 희소 복원 방법

3.1 그룹행렬 설계

먼저 그룹행렬 A 의 생성을 살펴보자. 그룹검사 시스템의 다양한 활용 분야의 특성에 따라 0이 아닌 원소가 희소한 그룹행렬 생성이 효율적이다. 그룹행렬 생성 시에 주의할 점은 0이 아닌 원소가 너무 많지 않도록 설계한다. 또한 하나의 샘플이 많은 그룹에 포함하여 그룹검사를 수행되지 않도록 하여야 하며, 하나의 그룹에 많은 샘플들을 포함하여 그룹검사를 하지 않도록 한다. 이를 해결한 가장 간단한 접근 방식은 랜덤 베르누이 행렬을 생성하여 활용하는 것이다. 이 행렬은 정규 분포의 랜덤 행렬과 함께 NSC(null space condition)과 RIP(restricted isometry property) 속성과 같은 조건을 충족한다[12]. RIP는 원래 신호를 높은 확률로 복원할 수 있음을 보장한다.

본 논문에서는 랜덤 베르누이 행렬 대신 LDPC 코드(Low-density parity-check code) [13] 방법을 이용한 그룹행렬을 사용한다. LDPC 코드는 1962년 Gallager가 처음 제안한 방법으로 새논의 정보 이론 값에 근접한 성능을 보여준다[14]. 본 논문에서 이용한 LDPC 코드 생성 방법은 Gallager가 최초로 제안한 방법을 따른다. 그룹행렬 생성은 각 행과 각 열의 1의 수가 일정하도록 하는 것이다. 예를 들어, 첫 행의 1의 수와 다른 행들의 1의 수가 같다는 것이고, 반면 첫 열에서의 1의 수와 다른 열에서의 1의 수가 같다는 의미이다. 다음의 그룹행렬은 열의 1의 수는 2이고, 행의 1의 수는 3인 경우이다.

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

논문 [15]에서는 각 열 1의 수 L 개로 일정한 그룹행렬이 RIP 조건을 만족함을 보여 주었다.

정리 2 [15]: 모든 열이 1의 수가 L 개로 일정하도록 생성된 그룹행렬 A 는 임의의 $\delta \in (0,1)$ 에 대해 높은 확률로 다음의 RIP 조건을 만족한다.

$$(1-\delta) \|X\|_2^2 \leq \|AX\|_2^2 \leq (1+\delta) \|X\|_2^2$$

3.2 희소 복원 방법

본 절은 제안하는 그룹검사 문제에 대해 신호를 복원하기 위한 희소 신호 복원 방법을 소개한다. 먼저 m 개로 구성된 신호 Y 로부터 복원하고자 하는 신호 X 를 찾는 방법은 다양하다. 첫 번째로 생각할 수 있는 것은 직접 대입법(brute force) 접근을 통해 가장 희소한 벡터 X 를 찾는다. 이 방법을 L_0 -norm 해법이라 불린다. 임의의 오차 ϵ 보다 작도록 하는 가장 희소 벡터를 찾는 것이다. 즉, 수식으로 표현하면 다음과 같다.

$$\min_X \|X\|_0 \text{ subject to } \|AX - Y\|_2 \leq \epsilon$$

그러나 L_0 -norm 방법은 모든 가능한 X 를 조합하여 문제를 해결하기 때문에 NP-hard가 된다. 2006년에 발표된 두 논문 [6]과 [12]에서 그룹행렬이 RIP 조건을 만족한다면 convex relaxation 기법을 활용하여 정확한 신호 X 를 찾을 수 있음을 밝혔다. 이 방법을 L_1 -norm 최소화 방법으로 불린다. Chen과 Donoho는 L_1 -norm 방법을 변형하여 다음의 최적화 방법을 제안한다[16].

$$\min_X \|AX - Y\|_2^2 + \lambda \|X\|_1$$

결국 기존의 L_1 -norm 최소화 방법과 위의 최적화 방법은 적절한 ϵ 과 λ 선택에 대해서 동일한 문제를 푸는 것과 같다. 본 논문에서는 X 의 원소가 음이 아닌 값의 형태를 갖는 점에서 다르다. 우리는 위의 최적화 방법을 좀 더 개선하여 그룹검사서 희소 신호를 복

원하는 문제로 다룬다. 다시 정리하면, 우리는 다음과 같은 최적화 문제를 고려하여 예측 신호 \hat{X} 를 찾는다.

$$\hat{X} = \arg \min_X \|AX - Y\|_2^2 + \lambda \|X\|_1, X \geq 0 \quad (7)$$

위의 최적화에서 \hat{X} 의 원소는 0과 1의 값이 아니기 때문에 우리는 임계값 $\tau = 0.5$ 을 설정하여 이 보다 크면 1이고 그렇지 않으면 0의 값을 갖는 것으로 설정한다.

3.3 성능 평가 및 논의

본 절에서는 앞서 설명한 기존의 그룹검사 알고리즘에 대해 컴퓨터 시뮬레이션을 통해 성능 평가 결과를 보여준다. 시뮬레이션을 위해 우리는 총 샘플의 크기를 $N = 1000$ 으로 설정하였다. 현재(2021년 11월 기준) 코로나19 발병률[17]이 각 나라별 또는 지역별, 연령대별로 다르지만 평균적으로 그 값이 1~5% 수준임을 고려하여 그룹검사에 대해 결함 샘플을 생성하여 시뮬레이션을 수행하였다. 그래서 본 시뮬레이션에서는 $K = 10$ (발병률 1%), $K = 30$ (발병률 3%), $K = 50$ (발병률 5%) 세 가지 설정조건을 적용한다. 알고리즘의 성능 평가 기준은 위음성이나 위양성이 나올 확률을 제외하여 정확하게 샘플의 상태(정상 또는 결함) 찾는지에 대한 성공 확률(successful probability)을 정의하여 사용한다. 실제 2021년 11월 기준, 코로나19 검사에서 95% 신뢰도 기준으로 위음성이 나올 확률이 0.1%에서 4.5% 수준이며, 위양성은 0%에서 0.3% 수준으로 보고되고 있다[18]. 본 논문에서는 수학적 모델이기 때문에 그룹검사에서 발생할 수 있는 다른 에러를 배제하고 도출된 결과임을 알아야 한다. 성공 확률 P_s 는 다음과 같다.

$$P_s = 1 - P(\text{위음성}) - P(\text{위양성}) \quad (8)$$

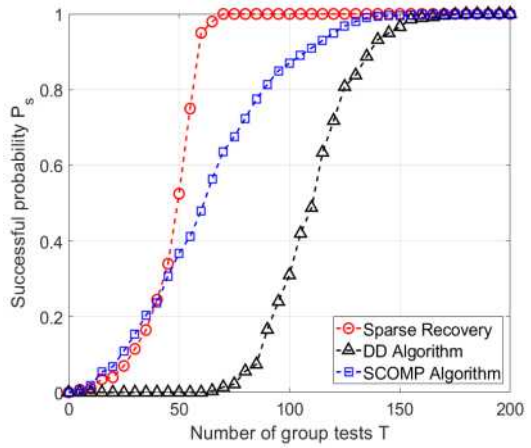


그림 1. 시뮬레이션을 통한 그룹검사의 성공 확률: $N = 1000, K = 10$.

Fig. 1. Successful probability on group testing using numerical simulation: $N = 1000, K = 10$.

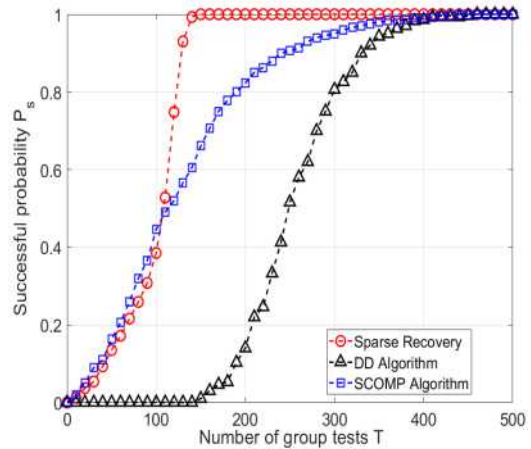


그림 2. 시뮬레이션을 통한 그룹검사의 성공 확률: $N = 1000, K = 30$.

Fig. 2. Successful probability on group testing using numerical simulation: $N = 1000, K = 30$.

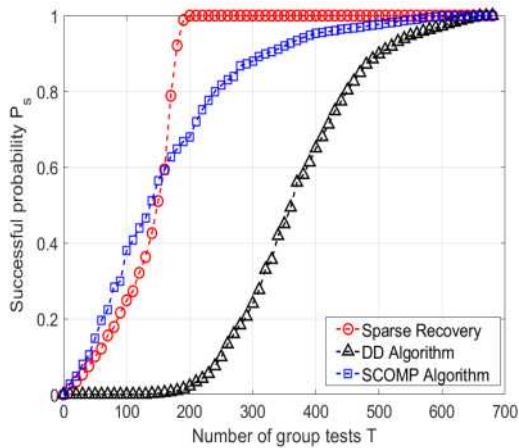


그림 3. 시뮬레이션을 통한 그룹검사의 성공 확률:

$N=1000, K=50$.

Fig. 3. Successful probability on group testing using numerical simulation: $N=1000, K=50$.

그림 1은 발병률이 1% 조건에서의 시뮬레이션 결과를 보여준다. 제안하는 희소 신호 복원 방법이 다른 알고리즘보다 더 우수한 성능을 보여준다. 특히 성공 확률이 절반 이상으로 올라갈수록 그룹검사 수를 조금만 더 수행하여 모든 샘플들의 상태를 더 빨리 찾을 수 있는 결과를 얻었다. 결국 희소 신호 복원 방법은 성공 확률이 0.5 근처에서 임계치로 작용하여 더 급격하게 성공 확률을 높이는 결과를 보여준다. 이와 같은 현상은 발병률 3%와 5%를 대상으로 한 $K=30$ 과 $K=50$ 조건에서 시뮬레이션을 수행한 그림 2와 3에서도 같은 모습을 보여준다.

그리고 발병률이 낮을 때보다 높을 때 제안하는 희소 신호 복원 방법이 더 우수한 결과를 얻었다. 다시 말하면, 성공 확률이 1에 가까울수록 SCOMP 알고리즘은 천천히 접근한 반면, 희소 신호 복원 방법은 더 빨리 다가갔음을 확인하였다. 그러나 그룹검사 수가 충분치 않고 적을 때는 희소 신호 복원 방법보다 SCOMP 알고리즘이 근소하게 더 좋은 결과를 보여 주었다. 그럼에도 불구하고 실제 그룹검사를 활용할 때는 모든 결함 샘플을 예리없이 찾을 수 있는지 여부가 더 중요하게 고려될 수 있으므로 낮은 성공 확률 영역에서의 성능 열화 문제는 큰 영향이 없을 것으로 간주된

다.

본 논문에서 제안하는 희소 신호 복원 방법을 이용하면 발병률 1%에서는 검사 수 70 이상, 발병률 3%에서 검사 수 150 이상, 발병률 5%에서 200 이상이면 성공 확률 0.99 이상으로 결함 샘플을 찾을 수 있음을 말한다. 시뮬레이션에서 총 1000개의 샘플을 대상으로 개별검사를 수행한 횟수보다 큰 검사 수의 감소를 가져왔다. 또한 제안하는 알고리즘은 기존의 DD와 SCOMP 복원 방법보다 복원 성능이 우수하다. 특히 결함 샘플을 모두 찾고자 하는 그룹검사 시스템에서는 다른 알고리즘에 비해 월등한 성능을 보여 주었다. 또한 결함 샘플이 적을 때보다는 많을 때 성능의 우수성을 더욱 두드러졌다.

4. 결론

본 논문에서는 그룹검사 문제를 살펴보았다. 최근 코로나19 확산과 피해를 최소화하기 위한 방법 중 가장 중요한 것은 감염자를 조기에 찾아내는 것이다. 이를 위해 그룹검사 방법을 통한 신속하고 정확하게 코로나19 감염자를 찾는 방법이 대두되고 있다. 본 논문에서는 기존에 발표된 여러 그룹검사 알고리즘들에 대해 동작원리를 살펴보았다. 그리고 압축센싱에서 제안한 희소 신호 복원 방법을 이용하여 그룹검사의 해법을 제시하였다. 압축센싱과 그룹검사는 연산 방법에서 차이가 있지만 희소 신호를 찾는다는 점에서는 유사하다. 시뮬레이션을 통해 기존의 알고리즘들과의 성능을 비교 평가하였다. 기존 알고리즘들과 비교하여 결함 샘플을 모두 정확히 찾고자 하는 그룹검사 시스템에서는 제안하는 희소 신호 복원 방법이 월등히 높은 성능을 보여 주었다. 또한 결함 샘플이 적을 때보다는 많을 때 그 성능의 우수성을 더욱 두드러졌다. 본 연구를 통해 얻은 연구결과를 활용하여 코로나19의 감염확산을 조기에 차단할 수 있는 하나의 방법이 될 것으로 기대된다.

REFERENCES

[1] R. Dorfman, "The Detection of Defective Members of Large Populations," *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 436-440, Dec. 1943.

[2] C.M. Verdun et al., "Group Testing for SARS-CoV-2 Allows for Up to 10-Fold Efficiency Increase Across Realistic Scenarios and Testing Strategies," *Frontiers in Public Health*, 9:583377, Aug. 2021.

[3] L. Mutesa, et al., "A pooled testing strategy for identifying SARS-CoV-2 at low prevalence," *Nature* 589, pp. 276-280, Oct. 2020.

[4] J.-T. Seong, "Group Testing-Based Robust Algorithm for Diagnosis of COVID-19" *Diagnostics* 2020, 10(6), 396, Jun. 2020.

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

[6] D.L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

[7] 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.

[8] 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.

[9] L. Baldassini, O. Johnson, M. Aldridge, "The capacity of adaptive group testing," *IEEE International Symposium on Information Theory*, pp. 2676-2680, Oct. 2013.

[10] T.M. Cover and J.A. Thomas, Elements of Information Theory, *Wiley*, Second Edition, 2006.

[11] J.-T. Seong, "A Lower Bound for Performance of Group Testing Problems," *Journal of Korea Institute of Information, Electronics, and Communication Technology*, vol. 11, no. 5. pp. 572-578, Oct. 2018.

[12] E.J. Candes, J. Romberg, T. Tao, "Robust

uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[13] R.G. Gallager, "Low density parity check codes," *IRE Transactions on Information Theory*, vol. IT-8, no. 1, pp. 21- 28, Jan. 1962.

[14] D.J.C. MacKay and R.M. Neal, "Near Shannon Limit Performance of Low Density Parity Check Codes," *Electronics Letters*, vol. 32, pp. 1645-1655. Jul. 1996.

[15] O. Johnson, M. Aldridge, J. Scarlett, "Performance of group testing algorithms with near constant tests per item," *IEEE Transactions on Information Theory*, vol. 65, no. 2, pp. 707-723, Feb. 2019.

[16] S. Chen S and D.L. Donoho, "Basis pursuit," *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers*, volume 1, pp. 41-44, 1994.

[17] Accessed: <https://www.worldometers.info/coronavirus/>

[18] D.A. Mistry et al., "A systematic review of the sensitivity and specificity of lateral flow devices in the detection of SARS-CoV-2", *BMC Infectious Diseases*, 21:828, 2021.

저자약력

정진택(Jin-Taek Seong) [정회원]



- 2014년: 광주과학기술원 정보통신공학과 (공학박사)
- 2008년 ~ 2010년: LG전자
- 2014년 ~ 2016년: 대구경북첨단의료산업진흥재단
- 2016년 ~ 2017년: 방위사업청
- 2018년 3월 ~ 현재: 목포대학교 융합소프트웨어학과 조교수

<관심분야> 정보이론, 압축센싱, 통신이론, 딥러닝