

## A Study on Monitoring System for Infectious Diseases Using Temporal Model with Application to Seoul Metropolitan Area

Hye Inn Song<sup>1#</sup>, Tae Young Heo<sup>2+</sup>, Yong Un Ban<sup>3</sup>

<sup>1</sup> Jeju Research Institute, 253 Ayeon-ro, Jeju Island, Korea

<sup>2</sup> Department of Information and Statistics, Chungbuk National University, 1, Chungdae-ro, Seowon-gu, Cheongju-si, Chungcheongbuk-do, Korea

<sup>3</sup> Department of Urban Engineering, Chungbuk National University, 1, Chungdae-ro, Seowon-gu, Cheongju-si, Chungcheongbuk-do, Korea

### Abstract

In this study, we proposed a systematic algorithm to monitor and detect outbreak of infectious disease and applied it to mumps data. The Farrington algorithm used in this study is a model that reflects temporal trend, seasonality, and over-dispersion of the data. It provides a reference value to judge outbreak by comparing it with observation. If observation at some time point is bigger than reference value, that time point is then assigned to be outbreak time. The application of the Farrington algorithm to the mumps data of Seoul's autonomous regions showed correct detection at increase points, and there are many outbreaks especially in the majority of autonomous districts from 2013 to the first half of 2014. The algorithm proposed in this study could provide more systematic and accurate information on disease surveillance and help the government to cope with outbreaks quickly and appropriately when applied to the monitoring system of various diseases.

**Key words:** outbreak, monitoring infection, Farrington algorithm, mumps

### 1. 서론

여러 가지 병원체에 의해 감염되어 발병하는 질환을 의미하는 감염병은 지구온난화를 비롯한 전 세계적인 기후 변화와 교통의 발달 및 생활 방식의 변화로 인해 신종 및 변종 감염병의 발생 빈도가 전 세계적으로 급속히 증가함에 따라 인류 건강과 경제를 위협하는 주요한 사안으로 급부상하고 있으며 전 세계적인 대유행으로의 폭발적인 확산 가능성이 매우 높아지고 있다. 또한

치사율이 높은 다양한 감염병들이 국내에 유입됨에 따라 이러한 감염병에 대한 과학적 감시체계 및 대응에 대한 연구가 필요한 실정이나 이러한 감염병에 대한 체계적인 데이터 관리 및 활용이 미흡한 실정이다.

감염병에 대한 대책은 이미 감염된 환자의 치료와 예방에 대한 체계와 함께 감염병의 확산을 막기 위해 감염병 급증을 초기에 탐지하는 것도 포함되고 있다. 감염병 감시 체계의 목적은 알려져 있지 않은 시간에서 발생한 프로세스의 변화, 즉 정상에서 벗어난 일탈을 탐지

<sup>#</sup> The 1st author: Hye Inn Song, Tel. +82-64-726-7401, e-mail. [shinn224@jri.re.kr](mailto:shinn224@jri.re.kr)

<sup>+</sup> Corresponding author: Tae Young Heo, Tel. +82-43-261-3741, e-mail. [theo@chungbuk.ac.kr](mailto:theo@chungbuk.ac.kr)

하는 것으로 국민 건강 증진을 위한 감염병 위기관리를 위한 질병 감시체계는 감염병의 관리 및 탐지에 매우 중요한 역할을 하고 있다.

최근 들어 생물학적 테러, 중증 급성 호흡기 증후군(SARS), 조류독감(Swine flue) 등과 같은 새로운 병원균의 지속적인 출현과 감염병 급증으로 국민 건강에 위협이 증가하고 있어 이러한 공공보건 문제를 해결하기 위해 통계적 감시시스템에 대한 많은 관심이 이루어졌다. 감염병 감시 체계는 국민의 건강 위협을 감소시킬 수 있으며, 미래에 어떤 일이 발생할지 알 수 있게 도움을 주며, 이러한 감염병의 위협을 관리하고 대응하기 위한 위기관리를 수립하는 데 활용할 수 있다.

감염병 급증에 대한 조짐을 보다 정확히 판단할 수 있는 감시체계의 구축이 필요하며, 이러한 감시체계는 축적된 감염병 발생건수를 이용하여 특정 시점의 급증 기준점을 제시할 수 있다.

감염병의 급증에 대한 조기 탐지 알고리즘은 기준값을 제시하고 이를 비교하는 방법들과 통계적 공정관리에 기인한 방법으로 구분된다(Farrington, *et. al.*, 1996; Rossi, *et. al.*, 1999; Höhle & Paul, 2008). 보다 구체적인 방법론으로는 회귀모형, 시계열 모형, 통계적 공정관리, 시공간 모형 그리고 다변량 기법 등과 같은 다양한 통계적 방법론이 제시되었으며 여러 연구들이 진행되었다(Farrington & Andrews, 2003; Sonesson & Bock, 2003; Lawson & Kleinman, 2005; Buckeridge, *et. al.*, 2005).

국내에서는 신종인플루엔자의 시공간 모니터링과 질병 확산 모형의 추정과 같은 연구가 수행되었으며(Lee & Hong, 2011; Seo & Choi, 2015), 신종인플루엔자의 시공간 모니터링의 경우 통계적 공정관리 방법을 이용한 감염병 감시체계를 제시하였다. 이와 달리 본 연구에서는 시계열 모형을 이용하여 추세 및 계절성을 반영한 보다 정확한 감염병 감시체계를 제시함에 의의가 있다.

본 연구에서는 특정 시점에 감염병의 급증의 여부를 결정할 수 있는 기준점을 제시하고, 해당 시점에서 발

생건수가 기준점보다 클 경우 급증이 발생하였음을 표시하는 알고리즘을 제시하였으며, 서울시 유행성이하선염 자료에 적용하여 급증을 판단하는데 유용함을 확인하였다. 본 연구를 통해 감염병 발생을 예측하고 조기에 유행을 억제할 수 있으며 감염병 유행 시 빠른 시일 내에 급증을 탐지하여 신속한 보건 정책 수립에 기여할 것이다.

## II. 급증 탐지 알고리즘

### 1. 급증 탐지 알고리즘 개념

본 연구의 목적은 과거의 시계열 자료를 기반으로 현재 시점에서의 질병의 급증을 보다 정확하게 판단하여 해당시점에 급증 여부를 판단하는 것이다. 본 연구에서 이용된 급증(outbreak) 탐지 알고리즘은 특정시점  $t_0$ 에 기준점을 계산하고  $t_0$ 시점의 감염병 발생건수가 기준값(threshold)보다 높을 경우 감염병 급증 발생을 의미한다. 이때 기준값은 참고값을 이용하여 알고리즘마다 각각 다른 방법을 사용하여 계산한다. 본 연구에서 사용한 급증 알고리즘에서는 종속변수가 관심 시점  $t_0$ 에서의 감염병 발생건수이며  $y_{t_0}$ 로 정의한다. 또한 해당 종속변수는 이산형 확률변수로 정의한다. 독립변수는 해당 시점  $t_0$ 와 시점의 주변 시점 및 과거의 해당시점과 과거 해당시점의 주변 시점이다. 여기서 주변시점이란 시점  $t_0$ 로부터 전후로  $n$ 일 또는  $n$ 주에 해당하는 시점이며, 과거 해당시점은  $k$ 년 전 시점  $t_0$ 에 해당한다. 자세한 내용은 2절에 설명하였다.

시점  $t_0$ 에서 급증 발생유무는  $OB(t_0)$ 로 표현되며 식 (1)과 같이 정의할 수 있다.

$$OB(t_0) = I(r(y_{t_0}) > g) \quad (1)$$

여기서  $r(y_{t_0})$ 는 시점  $t_0$ 에서의 예측값,  $g$ 는 기준값을 의미한다. 따라서  $OB(t_0)$ 는 지시함수로 감염병 급증여부를 0과 1로 구분하며 식 (1)을 이용하여 급증을

나타내는 시점들을 식 (2)와 같이 정의할 수 있다(Höhle & Mazick, 2010).

$$T_{OB} = \min \{t_0 \geq 1 : r(t_0) > g\} \quad (2)$$

급증 표시 알고리즘은 급증 여부를 판단하여 급증의 시점인  $T_{OB}$ 를 보다 정확한 방법으로 도출하는데 목적이 있다.

## 2. Farrington 알고리즘

본 논문에서는 급증 여부를 판단을 위해 급증 표시 알고리즘 중 가장 보편화된 알고리즘 Farrington 알고리즘을 이용하였으며, 정확도 향상을 위해 보다 개량된 알고리즘을 제안하였다. Farrington 알고리즘에 대한 간략한 설명은 다음과 같다. Farrington 알고리즘에서는 종속변수인 감염병 발생 건수가 가산자료이기에 과대산포를 보정하기 기준값을 도출하기 위한 통계 모형으로 퀴시 포아송 모형을 사용한다. 퀴시포아송 분포는 포아송 분포의 분산에 과분산 모수  $\lambda$ 를 곱한 형태이기 때문에 평균과 분산이 선형관계이며 과분산 모수가 1일 경우 평균과 분산이 같아져 일반적인 포아송 분포가 된다. 퀴시포아송 분포의 평균과 분산은 식 (3)과 같다.

$$E(Y_i) = \mu_i, \quad Var(Y_i) = \lambda_i \mu_i \quad (3)$$

여기서,  $Y_i$ 는 서울시  $i$ 번째 구에서의 감염자 수이고,  $\mu_i$ 는 평균 감염자수를 의미하며  $\lambda_i$ 는  $i$ 번째 구에서의 과분산 모수이며 퀴시포아송 모형의 모형식은 일반 포아송 모형의 식 (4)와 같다.

$$\log(\mu_i) = \alpha + \beta t_i \quad (4)$$

여기서  $t_i$ 는 시점을 의미한다.

본 모형에 사용하는 자료는 예측시점  $t_0$  전의 모든 시점이 아닌  $k$ 년도 전의  $t_0$ 시점과 동일한  $t_{0k}$ 시점과 그 주변 시점의 값을  $k$ 년도부터 현재년도까지 추출하여

사용한다. 이때 주변시점이란 해당 시점을 포함하여  $(2w + 1)$ 개의 시점을 의미하여,  $w$ 는 양수로 주 또는 일을 의미하며 알고리즘 적용 시  $k$ 와 함께 조절이 가능하다. 따라서 사용 자료개수는  $k$ 개의 년도와  $(2w + 1)$ 개의 주변시점으로 계산된  $k(2w + 1)$ 개다. 예를 들어 2015년의 7번째 주의 감염자수를 예측하고자 할 경우  $k$ 를 5,  $w$ 를 4라 하면 2015년의 3번째 주부터 6번째 주, 2014년의 3번째 주부터 11번째 주와 같이 2010년까지의 자료를 기준값으로 이용하는 것이며, 이는 모형에 계절성을 반영하게 한다(Höhle & Mazick, 2010).

Farrington 알고리즘에 사용되는 기준값 계산을 위해 사용하는 통계 모형은 식 (5)과 같다.

$$\log(\mu_t) = \alpha + \beta t \quad (5)$$

여기서,  $\mu_t, \lambda \mu_t$ 는 종속변수  $y_t$ 의 평균과 분산을 의미한다. 분산 내  $\lambda$ 는 과분산(overdispersion) 모수이며  $\alpha, \beta, \lambda > 0$ 이다.

Farrington 알고리즘은 식 (5)와 같이 추세를 반영하며, 시점  $t_0$ 에서의 신뢰구간 상한값을 기준값으로 사용하며 식 (6)과 같으며 관측된  $y_{t_0}$ 의 값이 기준값보다 클 경우 급증이 존재한다고 판단한다.

$$U_{t_0} = \hat{\mu}_{t_0} + z_{1-\alpha} \sqrt{Var(y_{t_0} - \hat{\mu}_{t_0})} \quad (6)$$

본 연구에서는 Farrington 알고리즘을 개선한 알고리즘으로 기존의 통계모형에 영차 주기 스플라인(zero order periodic spline)을 반영하여 급증 탐지 성능을 보완하였으며 식 (7)과 같다.

$$\log(\mu_t) = \alpha + \beta t + f(t) \quad (7)$$

여기서  $f(t)$ 이 영차 주기 스플라인으로 11개의 노드를 이용하였을 때 계절성을 무시하지 않는 것으로 알려져 있다(Noufaily, et. al., 2013).

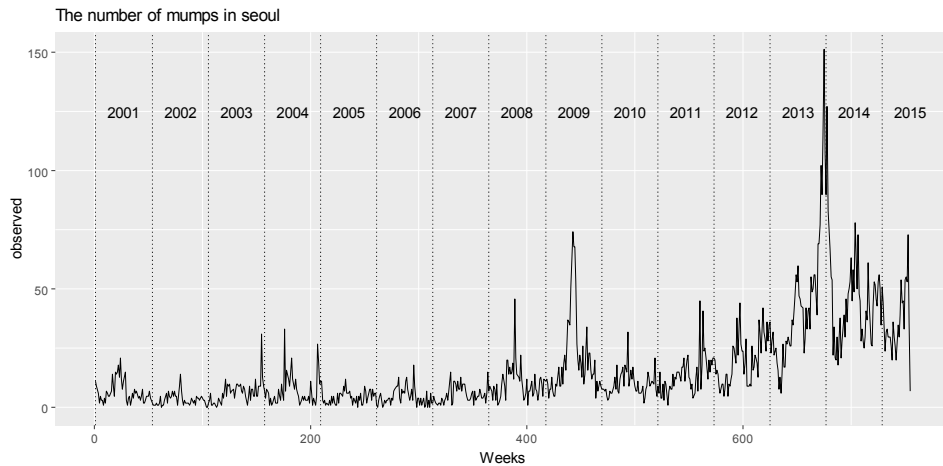


Figure 1. Time series plot of weekly mumps data in Seoul

### III. 자료 분석

#### 1. 자료설명

본 연구에서는 질병관리본부 감염병웹통계시스템<sup>1)</sup>에서 제공되는 서울시 각 행정구별 2001년 첫째 주부터 2015년 27번째 주까지의 총 755주이고, 해당 주별 유행성이하선염 발생건수 자료를 이용하였다. 시간에 따른 유행성이하선염 발생건수에 대한 시계열그림은 <Figure 1>과 같다. <Figure 1>을 보면 시간이 경과함에 따라 유행성이하선염의 발생건수가 증가하는 추세가 있으며 특히 2009년 및 2013년에서 2014년 사이에는 발생건수가 크게 증가했음을 알 수 있다. 또한 2013년과 2014년의 사이에 급증한 것을 제외하고 각 년도의 중반기에 발생건수가 다른 시기보다 증가한 것으로 나타남에 따라 해당 시계열 자료의 계절성이 존재함을 의심할 수 있다.

<Table 1>은 서울시 유행성이하선염 자료의 행정구별 기초통계량 및 쿼시포아송 분포에 기반을 둔 과분산 모수를 제시하였으며, 해당 과분산 모수는 1이상일 경우 과분산임을 의미한다. <Table 1>에서와 같이 과분산 모수가 가장 작은 구가 1,2032로 중구이고, 가장 높은 구가 11,0180으로 강남구이다. 따라서 모든 지역의 과분산모수가 1 이상으로 과분산이 존재한다고 할 수 있으며 해당 자료를 통합한 서울시의 경우도 과분산 모수

값이 21,0427로 과분산이 존재함을 알 수 있다. 이러한 과분산은 급증 탐지 알고리즘 적용 시 과분산 보정이 반영된 알고리즘을 사용해야하며, 본 연구에서 사용된 유행성이하선염 자료의 경우 시간적 추세, 계절성과 과분산이 존재함에 따라 Farrington 알고리즘의 적용이 적절하다고 할 수 있다.

### IV. 분석 결과

본 연구에서는 개선된 Farrington 알고리즘을 통해 서울시 25개 행정구에 대한 구별 분석을 실시하였고, 거시적인 해석을 위해 해당 구의 발생건수를 통합한 서울시 전체를 대상으로도 분석을 실시하였다. <Figure 2>와 <Figure 3>은 서울시 각 행정구별 인구수와 감염자수 사이에 관계를 확인하기 위해 2001년부터 2015년까지의 평균인구와 평균유행성이하선염 발생건수를 나타낸 지도이다. <Figure 2>와 <Figure 3>를 비교하면 감염자수가 가장 많은 지역과 인구수가 많은 지역이 비교적 유사한 것을 확인할 수 있으며, 이는 해당 행정구의 인구수가 많기 때문에 해당 지역의 발생건수도 많이 도출된 것으로 판단할 수 있다. 따라서 본 연구에서는 각 행정구별 인구수를 반영한 상쇄값(offset)을 식 (5)의 알고리즘에 포함하였다.

1) 질병관리본부 감염병웹통계시스템 : <https://is.cdc.go.kr/dstat/index.jsp>

Table 1. The basic statistics and overdispersion value by districts in Seoul

Region	Sum of Occurrences	Mean	SD	Over Dispersion
Jongro-gu	108	0,1430	0,4700	1,5442
Jung-gu	88	0,1166	0,3745	1,2032
Yongsan-gu	212	0,2808	0,8744	2,7228
Seongdon-gu	345	0,4570	1,0799	2,5522
Kwangjin-gu	496	0,6570	1,7481	4,6518
Dongdaemun-gu	703	0,9311	2,9421	9,2967
Jungrang-gu	522	0,6914	1,3297	2,5572
Seoungbuk-gu	503	0,6662	1,4891	3,3282
Kangbuk-gu	385	0,5099	1,3659	3,6586
Dobong-gu	444	0,5881	1,8364	5,7350
Nowon-gu	667	0,8834	2,0046	4,5484
Eunpyeong-gu	385	0,5099	1,1782	2,7222
Seodaemun-gu	201	0,2662	0,5798	1,2628
Mapo-gu	395	0,5232	0,9443	1,7044
Yangcheon-gu	911	1,2066	2,1569	3,8556
Gangseo-gu	479	0,6344	1,3612	2,9206
Guro-gu	661	0,8755	1,6296	3,0332
Keumchen-gu	348	0,4609	1,3339	3,8603
Yongdeungpo-gu	332	0,4397	1,3450	4,1139
Donjak-gu	565	0,7483	1,2763	2,1767
Seocho-gu	312	0,4132	0,9235	2,0639
Kangnam-gu	662	0,8768	3,1080	11,0180
Songpa-gu	771	1,0212	1,6171	2,5607
Kangdong-gu	514	0,6808	1,2199	2,1859
Kwanak-gu	513	0,6795	1,1519	1,9527
Seoul(Total)	11,522	15,2609	17,9200	21,0427

Seoul population map

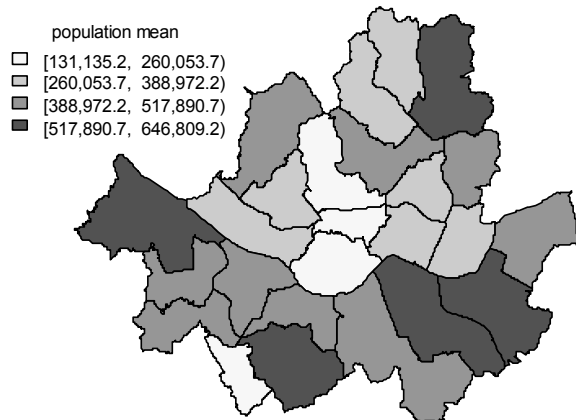


Figure 2. Map for average population of 25 districts of Seoul

Seoul mumps map

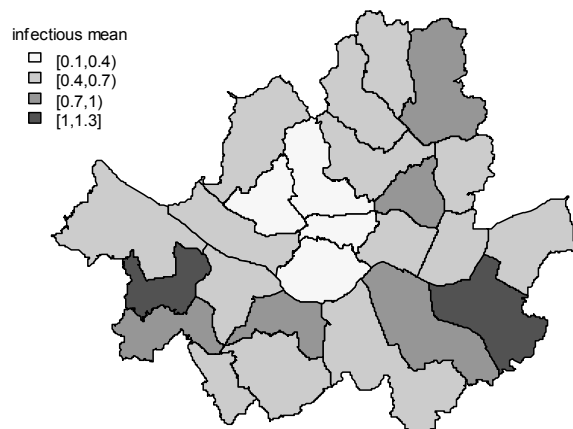


Figure 3. Map for average number of mumps occurrences of 25 districts of Seoul

### Surveillance using Farrington Algorithm

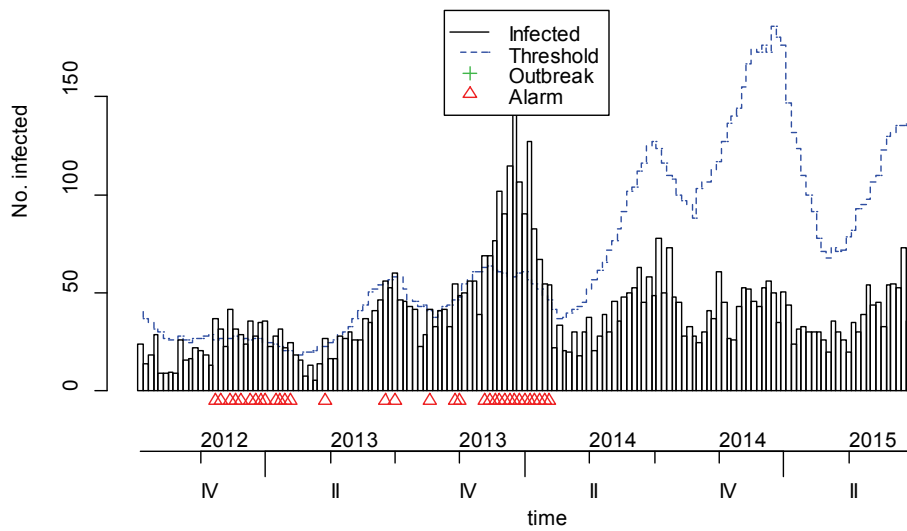


Figure 4. The result of farrington algorithm with Seoul(total) data

본 연구에서 사용한 Farrington 알고리즘은 감염병에 대해 보다 정확한 급증의 기준을 제시하여 감염병 급증에 대한 대응을 가능하게 한다. 본 논문에서는 서울시 전체 및 각 구별의 가장 최근 주인 755주부터, 155주 전인 600주까지의 급증 여부를 확인해보았다. 이때 각 주의 참고값으로 해당 주를 기준으로 4년 전부터 해당 주와 해당 주의 앞뒤 3주씩의 자료를 사용하였다.

첫 번째로 서울시 전체에 대한 Farrington 알고리즘 적용결과를 <Figure 4>에 제시하였다. 그림에서 막대는 각 주의 유행성이하선염 발생건수이고, 파란색 선은 Farrington 알고리즘을 통해 계산한 기준값이며, 붉은색 삼각형은 급증이 발생한 지점을 의미한다. 여기서 범례에 제시된 녹색 '+' 표시는 실제 급증이라고 판명된 경우 그림에 포함되나 본 연구에서 사용한 자료를 실제로 급증이 발생하였다고 판명된 경우가 존재하지 않아 표시되지 않았다.

Farrington 알고리즘의 적용 결과 2012년에 유행성이하선염이 급증되었음이 표시되었고, 2013년에 유행성이하선염 발생건수가 매우 증가하였을 때도 해당 시점이 급증임이 잘 표시되었다. 또한 2013년 이후에 발생건수가 감소하였으나 2012년도 발생건수보다는 크

게 나타났음에도 급증표시가 되지 않은 것을 확인할 수 있는데, 이는 유행성이하선염 데이터의 추세가 반영되어 해당 발생건수가 급증이 아닌 2013년에 비해 감소한 경우임을 나타낸다. 결과적으로 서울시는 2012년부터 2014년 전반기까지 유행성이하선염이 급증하였고 2014년 후반기부터 급증으로 판단되는 발생건수는 나타나지 않았다고 할 수 있다.

두 번째로 서울시의 각 행정구별 Farrington 알고리즘의 적용 결과를 <Figure 5>에 제시하였다. 그림을 보면 급증이 발견된 대부분의 지역이 광진구, 노원구, 양천구, 강동구를 제외하고 2013년부터 2014년도에 급증이 밀집되어 있는 것으로 나타났다. 광진구, 노원구, 양천구, 강동구의 경우 다른 년도에도 급증 발생이 존재하였으나 특히 2012년도에 급증이 밀집되어 있다. 또한 2013년에 강남구와 동대문구에 30건 이상의 큰 발생건수가 집계되었고 해당 시점이 급증 발생 시점으로 판명되었는데, 해당 구 주변의 구도 급증 발생 시점이 유사하여 공간상관성이 존재하는지를 확인하기 위해 동대문구와 강남구에서 가장 발생건수가 많은 시점에서 급증 발생 유무 지도를 제작하였다. 지도는 <Figure 6>과 <Figure 7>에 제시하였다.

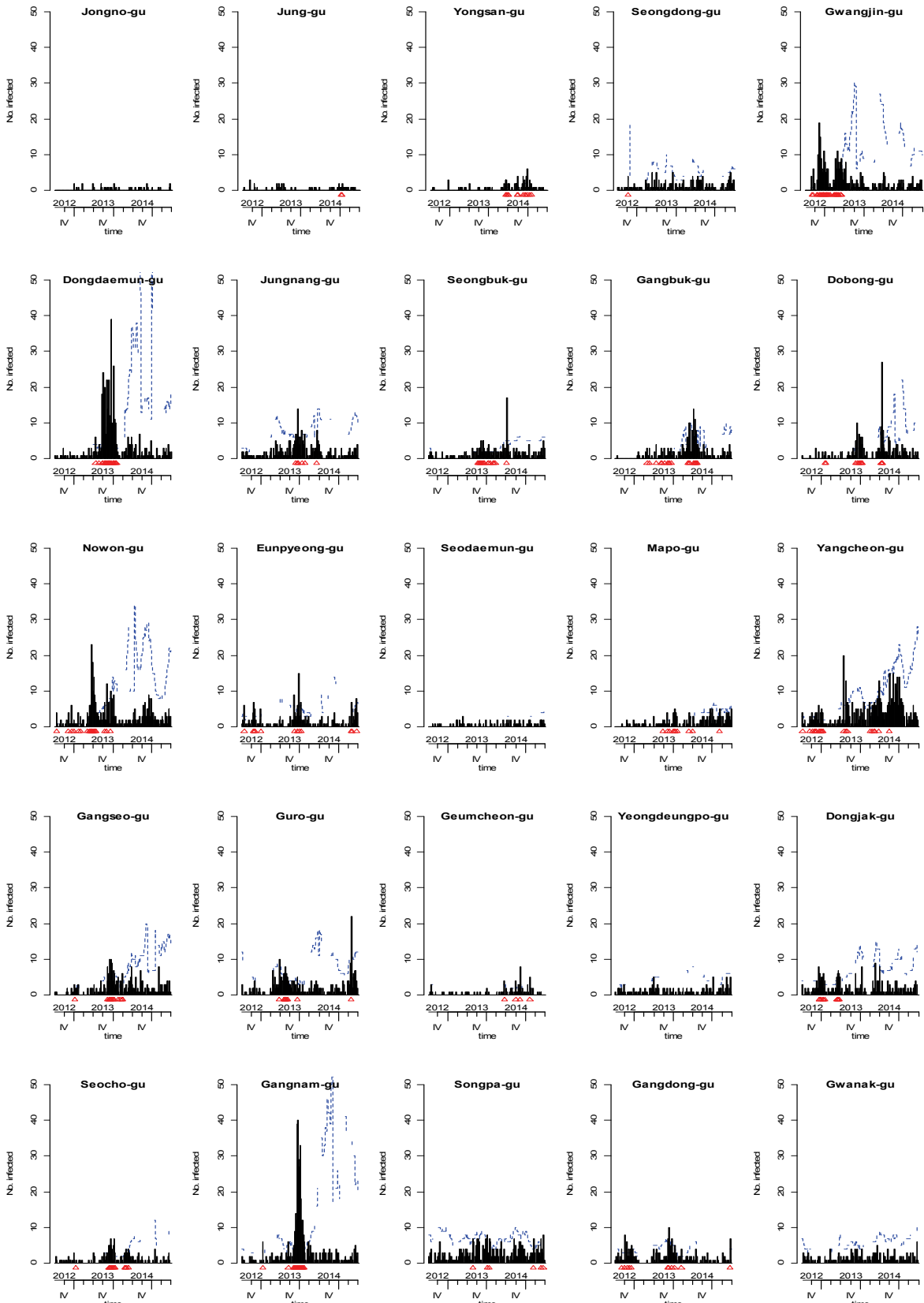


Figure 5. The result of farrington algorithm with 25 districts of Seoul data

Seoul mumps alarm map

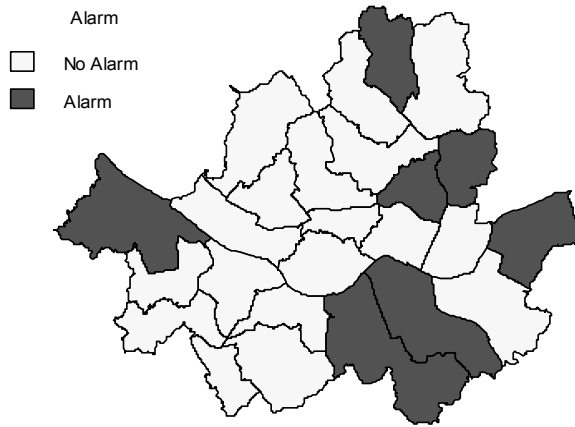


Figure 6. The result of alam map on 675th week data by proposed algorithm

〈Figure 6〉은 강남구와 동대문구에서 가장 많이 불거리가 발생한 주인 675번째 주의 서울시의 급증 유무를 지도에 나타낸 것이다. 그림을 보면 강남구와 서초구, 동대문구와 중랑구가 인접해있어 공간상관성이 존재함을 의심할 수 있다. 기간을 해당 주가 포함된 672번째 주부터 681번째 주까지 확대하여 모든 구에서의 급증 표시 건수를 지도에 표현한 결과를 〈Figure 7〉에 제시하였다. 해당 기간은 강남구에서 급증이 가장 많이 발생한 기간이다.

그림을 보면 강남구와 서초구, 도봉구와, 강서구, 동대문구에서 해당 기간 동안 급증이 가장 많이 발생하였고, 강서구와 동대문구에 인접해있는 구에서 3건에서 6건 사이의 급증이 발생한 것으로 나타났다. 이에 유행성이하선염의 급증이 공간상관성이 존재하며 이는 유행성이하선염이 감염병이기 때문인 것으로 판단된다. 알고리즘 적용 결과 감염병의 급증에 대한 판단을 위해 본 연구에서 사용된 Farrington 알고리즘을 적용할 경우 급증의 판단은 보다 정확하게 이루어질 수 있으며, 각 자치구 별 및 전체 지역의 급증 파악이 가능하므로 거시적 및 미시적인 분석이 가능함을 확인하였다. 그러나 해당 알고리즘은 공간상관성을 반영하지 않아 유행성이하선염이 공간상관성이 존재함이 의심됨에도 해당 영향을 반영하지 못하는 단점이 있다.

Seoul mumps alarm map

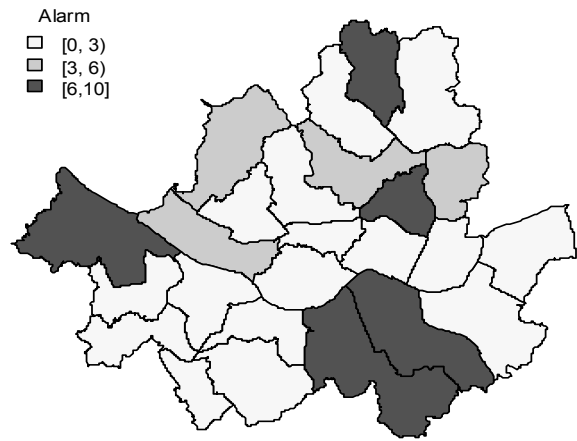


Figure 7. The result of alam map on 672th-682th week data by proposed algorithm

### V. 결론

조기 감염병 급증 탐지 알고리즘을 통해 질병의 감염 확산의 위험을 감소시키거나 제거할 수 있으며 해당 위험요인들에 대해 경고할 수 있다. 또한 감염병 확산에 대한 조기 진단 및 탐지를 통해 신속한 방역체계를 구축할 수 있으며 감염병 급증에 따른 백신개발 시기 결정, 선제적인 보건 정책 개발 등과 같이 감염병 조기 탐지는 매우 중요하다.

본 연구에서는 감염병의 급증에 대한 정확한 판단을 위해 Farrington 알고리즘을 제시하였으며, 서울시의 2001년부터 2015년도까지의 유행성이하선염 자료에 Farrington 알고리즘을 적용한 결과 감염병 급증에 대한 판단이 적절하게 이루어졌음을 확인하였다.

본 연구에서 제안한 알고리즘을 통한 감염병 급증 판단의 결과는 감염병의 모니터링 시스템에 적용될 경우 보다 정확한 상황 파악 및 대처가 가능하며, 다양한 감염병에 대한 발생건수의 데이터베이스가 축적됨에 따라 쉽게 적용이 가능할 것으로 판단된다.

또한 보건 및 질병 관련 모니터링을 위해 사용하는 자료로서 특정 감염병 발생건수 자료뿐만 아니라 국민건강보험공단 및 건강보험심사평가원에서 수집하는 국민건강보험자료를 가공하여 특정 질병의 처방 건수 또는 매일 자료가 생성되는 각 병원에서의 진료 자료 등을 모니



터링 한다면 보다 정확하게 빠르게 감염병의 급증에 대한 판단을 할 수 있다. 그러나 본 연구에서 제안한 통계 모형 기반의 감염병 급증 탐지 알고리즘은 과거에 발생한 감염병의 수를 이용하고 분석하는 기법으로 과거 발생자료의 누적 여부와 자료의 신뢰성이 매우 중요하다.

본 연구에서 사용한 Farrington 알고리즘은 쿼시포 아송 모형과 함께 영차 주기 스플라인(zero order periodic spline) 방법을 사용하여 감염병 자료의 급증을 판단하는데 보다 정확한 결과를 제시하나 공간상관성을 반영하지는 않는다는 한계가 있다, 따라서 본 연구에서는 서울시를 행정구로 나누어 Farrington 알고리즘을 적용할 때 공간상관성을 반영하여 모형이 적합 되는 것이 아니라 각 행정구별로 알고리즘이 개별 적용되어 급증을 판단하였다. 따라서 Farrington 알고리즘은 서울시 내 행정구별 공간상관성은 반영이 되지 않아 각 행정구별 시점에 따라서만 기준값을 계산하기 때문에 급증(outbreak)을 감지하는데 정확성 및 민감성이 해당 내용을 반영한 알고리즘에 비해 상대적으로 낮을 수 있다. 따라서 향후 연구로서 공간상관성이 반영과 함께 한 시점에 대한 기준값 제시가 아닌 기준값들의 누적을 이용한 알고리즘 등을 개발하여 보다 좋은 성능을 가진 감염병 판단기준 알고리즘을 제안하고자 한다.

## 감사의 글

이 논문은 2016년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2016S1A5B6914198).

## References

Buckeridge, D., H. Burkom, M. Campbell, W. Hogan, and A. Moore. 2005. Algorithms for Rapid Outbreak Detection: A Research Synthesis. *Journal of Biomedical Informatics*. 38(2): 99-113.

Farrington, C. and N. Andrews. 2003. Monitoring the Health of Populations. Oxford University Press. *Outbreak Detection: Application to Infectious Disease Surveillance*. 203-231.

Farrington, C., N. Andrews, A. Beale, and M. Catchpole. 1996. A Statistical Algorithm for the Early Detection of Outbreaks of Infectious Disease. *Journal of the Royal Statistical Society, Series A*. 159: 547-563.

Höhle, M. and A. Mazick. 2010. Aberration Detection in R Illustrated by Danish Mortality Monitoring. *Biosurveillance: Methods and Case Studies*. 215-237.

Höhle, M. and M. Paul. 2008. Count Data Regression Charts for the Monitoring of Surveillance Time Series. *Computational Statistics & Data Analysis*. 52(9): 4357-4368.

Lawson, A. and K. Kleinman. 2005. *Spatial and Syndromic Surveillance for Public Health*. Wiley.

Lee, Gyoung Ju and Il Young Hong. 2011. A Study on the Methodology of Spatial-temporal Monitoring to Establish a System of Beforehand Surveillance on Influenza. *Journal of the Korean Official Statistics*. 16(2): 22-37.

Noufaily, A., D. G. Enki, P. Farrington, P. Garthwaite, N. Andrews, and A. Charlett. 2013. An Improved Algorithm for Outbreak Detection in Multiple Surveillance Systems. *Statistics in Medicine*. 32(7): 1206-1222.

Rossi, G., L. Lampugnani, and M. Marchi. 1999. An Approximate CUSUM Procedure for Surveillance of Health Events. *Statistics in Medicine*. 18: 2111-2122.

Seo, Mi Ok and Bo Seung Choi. 2015. A Estimation Method for Stochastic Epidemic Model. *Journal of the Korean Data Analysis Society*. 17(3): 1247-1259.

Sonesson, C. and D. Bock. 2003. A Review and Discussion of Prospective Statistical Surveillance in Public Health. *Journal of the Royal Statistical Society, Series A*. 5-12.

## Korean References Translated from the English

서미옥, 최보승. 2015. 확률적 방법을 이용한 질병 확산 모형 추정 방법. *한국자료분석학회지*. 17(3): 1247-1259.

이경주, 홍일영. 2011. 신종 인플루엔자의 사전적 감시체계 구축을 위한 시공간 모니터링 방법론 활용에 관한 연구. *통계연구*. 16(2): 22-37.

Received: Sep. 22, 2017 / Revised: Nov. 3, 2017 / Accepted: Nov. 9, 2017

## 시간 모형을 이용한 감염병 감시 체계에 대한 연구

- 서울시를 중심으로 -

국문초록 본 연구는 전염병 감시에 있어 해당 시점에서의 급증 여부를 판단할 수 있는 체계적인 알고리즘을 제안하고자 급증 탐지 알고리즘을 서울시의 구별 유행성이하선염 자료에 적용하였다. 본 연구에서 사용된 Farrington 알고리즘은 자료의 시간적 추세 및 계절성, 과분산을 반영한 모형으로 급증을 판단할 수 있는 기준치를 제시하고, 전염병 자료의 각 시점의 발생건수와 기준치를 비교하여 기준치 이상의 발생건수가 집계된 시점을 급증 시점으로 판단한다. Farrington 알고리즘을 서울시의 구별 유행성이하선염 자료에 적용한 결과 눈에 띄게 발생건수가 증가한 시점에 급증이 알맞게 표시되었고, 특히 2013년부터 2014년 전반기까지에 대다수의 자치구에 급증이 많이 표시되었다. 본 연구에서 제안한 알고리즘은 향후 감염병의 감시에 있어 보다 체계적이고 정확한 정보를 제공할 것이며 해당 알고리즘이 여러 질병의 모니터링 시스템에 적용될 경우 적절하고 빠르게 급증에 대한 대처가 가능할 것으로 기대된다.

주제어 : 급증, 전염병감시체계, 알고리즘, 유행성이하선염

Profiles **Hye Inn Song** : She received the B.S and M.S. in Statistics from the Chungbuk National University. Now she is researcher in Jeju Research Institute. Her current research interests include public transportation modeling based on big data([shinn224@jri.re.kr](mailto:shinn224@jri.re.kr)).

**Tae Young Heo** : He received the B.S. in Statistics from the Chungbuk National University and the M.S. and Ph.D in Statistics from the North Carolina State University. Now he is a professor in department of Information & Statistics, Chungbuk National University. His current research interests include spatial modeling, transportation and environmental statistics based on various spatial methodologies([theo@cbnu.ac.kr](mailto:theo@cbnu.ac.kr)).

**Yong Un Ban** : He received his M.A. from Yonsei University and Ph.D. from University of Pennsylvania. He is a professor of the Department of Urban Engineering at Chungbuk National University. His research interests included sustainable development, urban spatial configuration, LID(Low Impact Development), EIP(Eco Industrial Park), climate change, and environmental justice. He has published 78 articles in journals and written 5 co-author books([byubyu@chungbuk.ac.kr](mailto:byubyu@chungbuk.ac.kr)).