

순간 재생산지수에 기반한 COVID-19 확진자 수 변화 예측 시스템 연구

안수진, 권민혜

송실대학교 브레인 및 기계 지능 연구실

asujin331@soongsil.ac.kr minhae@ssu.ac.kr

COVID-19 Forecast System Based on Instantaneous Reproduction Number

Sujin Ahn Minhae Kwon

Soongsil University

요 약

코로나 바이러스(SARS-CoV2; COVID-19)로 인해 대한민국은 거리 두기 정책이라는 비약물적 중재와 백신이라는 약물적 중재를 통합적으로 사용함으로써 방역의 효과를 극대화하고 있다. 팬데믹으로 인해 예측 불가능한 상황의 도래가 불가피해지면서 인적 및 물적 피해를 최소화할 수 있는 방역 정책을 수립하기 위해서는 감염병 확산 예측에 대한 연구가 필수적이다. 특히, 감염 재생산지수(reproduction number)에 대한 예측은 감염병 전파(transmission)의 변화 양상을 감지하여 방역 정책과 관련된 의사 결정시에 중요한 판단 기준을 제공할 수 있다. 본 논문은 순간 감염재생산지수를 기반으로 가까운 미래의 일일 확진자 수를 예측하는 시스템을 제안한다.

I. 서 론

COVID-19 유행 확산 예측은 미래의 확진자 수를 예측함으로써 향후 방역 정책의 집행과 의료 인력 배치 및 필요 병상 확보 등의 의사결정에 도움을 준다. 이러한 예측 연구는 위기 관리 프로세스 마련과 종합적인 팬데믹 대응책 결정 등에 효과적인 기준점을 제시한다. 본 연구에서는 감염병 전파 양상을 파악하는데 중요한 지표인 감염재생산지수에 대해 설명하고, 미래의 일일 확진자 수 예측을 위한 시스템을 제안한다.

수식(1)은 확진자 수의 시간에 따른 변화율로, $\frac{dI(t)}{dt} > 0$ 이면 확진자 수가 증가하고 있음을 의미한다.

확진자 수가 증가하는 상황을 가정하고, 수식(1)을 정리하면 SIR 모델에서의 기초 감염재생산지수(basic reproduction number; R_0)에 대해 표현할 수 있다. 여기서 R_0 는 감염자가 없는 인구 집단에 처음으로 감염자가 발생하였을 때 첫 감염자가 평균적으로 감염시킬 수 있는 2 차 감염자의 수를 나타낸 것이다[3]. 따라서, $S(t) = N$ 으로 표현 가능하다.

$$R_0 = \frac{\beta S(t)}{\gamma N} = \frac{\beta N}{\gamma N} = \frac{\beta}{\gamma} > 1 \quad (2)$$

수식(2)을 통해 R_0 가 1 보다 크면 감염병이 확산되는 조건을 만족시킴을 알 수 있다.

II. SIR 모델 및 감염재생산지수 연구

II.1 SIR 모델

감염병 전파에 관한 대표적 구획모델으로는 Susceptible-Infected model[1]과 Susceptible-Infectious-Recovered model(SIR 모델)이 있다. SIR 모델은 질병에 걸릴 수 있는 감수성자의 상태(susceptible; S)와 감염 상태(infectious; I), 회복 상태(recovered; R)라는 구획으로 구성된다. 감수성자의 상태에서는 감염 상태의 개체와의 접촉을 통해 감염 상태가 되며, 이후 회복되거나 사망하게 되면 감염 상태에서 회복 상태가 된다. t 시점에서 누적 감수성자의 수를 $S(t)$, 누적 감염자 수를 $I(t)$, 누적 회복자 수를 $R(t)$ 라고 한다. SIR 구획의 시간에 대한 변화율은 다음과 같이 나타낼 수 있다[2].

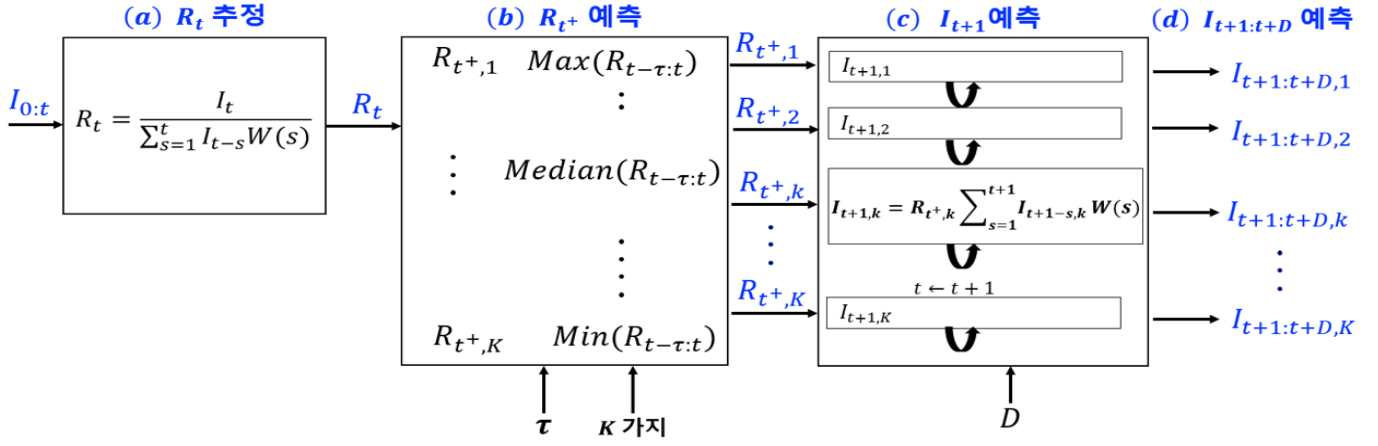
$$\begin{aligned} \frac{dS(t)}{dt} &= -\frac{\beta S(t)I(t)}{N} \\ \frac{dI(t)}{dt} &= \frac{\beta S(t)I(t)}{N} - \gamma I(t) \\ \frac{dR(t)}{dt} &= \gamma I(t) \end{aligned} \quad (1)$$

N 은 $S(t) + I(t) + R(t)$ 로, 전체 인구수로 정의된다. β 는 전파율(disease transmission rate), γ 는 회복률 또는 제거율(recovery or removal rate)이다.

II.2 감염재생산 지수

감염재생산지수는 한 명의 감염자가 평균적으로 감염시킬 수 있는 2 차 감염자의 수로 정의한다. 수식(2)와 같이 감염재생산지수가 1 보다 크면, 한 명의 감염자가 한 명 이상의 2 차 감염자를 발생시키며, 이는 감염병의 확산으로 이어진다. 감염재생산지수가 1 이면 풍토병(endemic)이 되고, 1 보다 작으면 감염병의 확산이 감소한다. 감염재생산지수에는 기초 감염재생산지수, 실질 감염재생산지수(effective reproduction number; R_e)가 있다. R_0 가 감염병 발생 초기 병의 감염력을 나타내는 지표라면, R_e 는 일정 시간이 지난 후 개인 방역과 사회적 정책 실시까지 고려한 감염재생산 지수를 의미한다[4].

실질 감염재생산지수 중 실시간 감염 재생산 지수(time-varying reproduction number)가 있다. 실시간 감염재생산지수는 순간 재생산지수(instantaneous reproduction number)와 케이스 재생산지수(case reproduction number) 두 가지 방법으로 표현된다. 순간 재생산지수는 특정 시간의 감염재생산지수를 추정하거나 정책, 행동, 외부적 요인으로 인한 전파를 추론하는 실시간 분석에 널리 사용한다.



[그림 1] 순간 재생산지수에 기반한 COVID-19 일일 확진자 수 예측 시스템

반면, 케이스 재생산지수는 서로 다른 시점에서 감염된 특정 개인의 감염재생산지수를 분석하는 후향적(retrospective) 연구에 사용한다[5]. 순간재생산지수에 기반한 예측 시스템은 실시간 의사 결정시에 유용하므로, 본 연구에서는 Cori[6]의 방법론을 적용한 예측 프로그램을 제시하였다. Cori의 방법론은 확진자 수와 감염성을 이용하여 t 시점의 순간 감염재생산지수 R_t 를 추정하는 것이다. 여기서, t 는 확산의 시작일로부터 t 번째 일로 추정한다.

$$R_t = \frac{I_t}{\sum_{s=1}^t I_{t-s} W(s)}, \quad t \geq 1$$

I_t 는 확산의 시작일로부터 t 번째 일의 확진자 수이고, I_0 는 확산의 시작일을 나타낸다. $W(s)$ 는 감염된 후 s 일이 지난 후인 현재 감염성(current infectiousness)으로 표현된다.

$$W(s) = \left[\frac{1}{\Gamma(\alpha)\theta^\alpha} \right] s^{\alpha-1} e^{-\frac{s}{\theta}} \quad (3)$$

여기서, α 는 모양 매개변수(shape parameter)이고, θ 는 크기 매개변수(scale parameter)로 $\Gamma(\alpha) = (\alpha-1)!$ 이다. 현재의 감염성 추정을 위해 감염자의 증상 발현일로부터 2차 피감염자의 증상 발현일까지의 기간을 나타내는 세대기(serial interval)를 이용한다. 세대기는 WHO-PAHO(Pan American Health Organization)에 의해 평균 4.8, 표준편차 2.3을 따르는 감마 분포를 활용한다.

III. 순간 감염재생산지수 기반 일일확진자수 예측 시스템 제안

본 예측 시스템은 t 를 기준으로 과거 τ 일 간의 순간 감염재생산지수에 대해 K 가지 적용 가능한 예측 방법을 고려하여 미래 D 일에 대한 확진자 수를 추정하고자 한다. 본 연구에서 제안하는 예측 프로세스는 그림 1에 설명되어 있다. $I_{0:t}$ 는 확산의 시작일로부터 t 번째 일까지의 확진자 수이다. 먼저 $I_{0:t}$ 를 입력값으로 받아 $W(s)$ 와 확진자 수를 이용하여 R_t 값 계산한다(그림 1(a)). COVID-19의 경우 $W(s)$ 는 수식 (3)을 따른다. 다른 감염병의 경우, $W(s)$ 를 해당 감염병의 특성에 맞게 변화시킬 수 있다. 이후 $t-\tau$ 일부터 t 일까지의 데이터를 이용하여 K 가지의 예측 방법을 이용하여 R_{t+} 를 예측한다(그림 1(b)). 여기서, R_{t+} 는 t 번째 일 이후의 미래 순간 감염재생산지수로, 본 시스템을 통해 예측된 값이다. K 개의 예측 방법 설정에 대해서는 현재 다양한 연구가 진행 중에 있다[7]. 가장 대표적인 방법으로는

t 로부터 최근 과거 τ 일 간의 순간 감염재생산지수 중 최댓값, 중앙값, 최솟값들을 뽑아내어 이를 R_{t+} 로 추정하는 방법이 있다. 마지막으로, 앞서 결정한 R_{t+} 와 $I_{0:t}$ 를 이용하여 일일 확진자 수인 I_{t+1} 를 추정한다(그림 1(c)). I_{t+1} 을 추정하기 위해 $I_t = R_t \sum_{s=1}^t I_{t-s} W(s)$ 을 이용하였다. 이 과정을 D 번 반복함으로써 미래 D 일에 대한 일일 확진자 수인 $I_{t+1:t+D}$ 를 예측할 수 있다(그림 1(d)).

IV. 결론

본 연구는 일일 확진자 수를 이용하여 순간 감염재생산지수를 추정하고, 미래의 확진자 수를 예측하는 시스템을 제안한다. 제안한 시스템은 상황에 따라 다양한 예측 방법론을 선택 가능하고, 다른 감염병에도 적용할 수 있다는 점에서 적용 범위가 넓다는 장점이 있다. 본 연구의 예측 시스템은 단기 예측을 통해 유행 확산 추이를 파악하고, 정책 마련, 의료 인력 배치 등 효율적인 대비를 위한 근거를 마련하고자 한다.

ACKNOWLEDGEMENT

이 논문은 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업(IITP-2021-2020-0-01602)과 방송통신산업기술개발사업(IITP-2021-0-00739), 그리고 한국연구재단(NRF-2020R1F1A1069182)의 지원을 받아 수행된 연구임.

참 고 문 헌

- [1] 안수진, 권민혜, “대한민국 COVID-19 확산과정에서의 SI 모델 기반 사회적 거리두기 정책효과 분석,” 통신정보 합동학술대회(JCCI), April 2021.
- [2] Fred Brauer, “Mathematical Models in Epidemiology,” Springer, 2019.
- [3] James Holland Jones, “Notes on R0,” California: Department of Anthropological Sciences, 2007.
- [4] 유명수, 김연주, 백수진, 권동혁, “감염재생산지수 개념 및 방역 정책에 따른 변화,” KDCA, 제 14 권, 6 호, 2021년 2월.
- [5] Gostic KM et.al, “Practical Considerations for Measuring the Effective Reproductive Number, Rt,” PLoS Computational Biology, 2020.
- [6] Anne Cori et.al, “A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics,” American Journal of Epidemiology, vol.178, 2013.
- [7] 권오규, 손우식, 심은하, 이창형, 이효정, 정은옥, 정일호, 최선화, 황형주, “수리모델링으로 분석한 코로나 19 유행예측,” 국가수리과학연구소, 2021년 5월.