# 스마트시티 연구소는 코로나 사태를 극복하기 위해 어떤 노력을 하고있을까

## 서론

봄은 많은 이들이 기다리는 계절입니다. 추위를 깨고 꽃이 피며 사람들의 마음을 설레이게 만듭니다. 그러나, 코로나 바이러스 때문에 많은 사람들이 화창한 날씨에도 불구하고 밖에 나가지 못하고 있습니다. 백신이 존재하지 않고 잠복기가 존재하기 때문에 누가 보균자인지 판단하기도 어렵습니다.

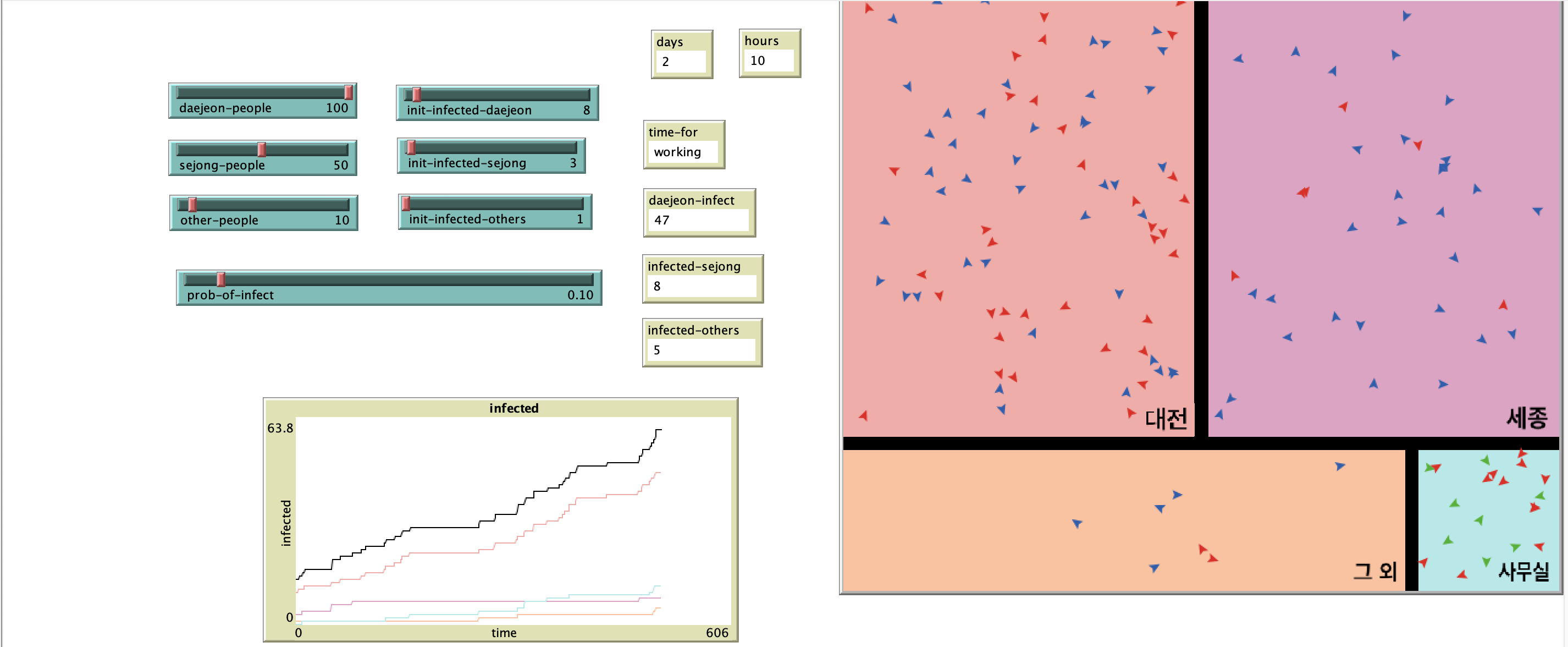
코로나-19 사태가 장기화 된다면 어려움을 겪는 사람들이 많아지게 됩니다. 이러한 문제에 맞서기 위해 각 분야에서 많은 전문가 분들이 노력해주고 계시며 다음소프트 스마트시티 연구소도 이 문제에 대해 고민해보았고, 어떤 방법으로 문제 해결을 위해 노력했는지 공유하고자 합니다.

## Netlogo를 이용한 간단한 실험

문제를 해결하기 위해서는 어떤 위험성을 가지고 있는지, 먼저 알아볼 필요가 있습니다. 어떤 전파 양상을 띄는지, 어느 시기에 얼마나 많은 사람들이 감염될 지를 알면 그에 적절한 대응도 가능할 것이라고 생각했습니다. 그러기 위해서는 현재 가지고 있는 정보를 취합해 **시뮬레이션**을 해보는 것이 가장 적절한 방법이라고 생각했습니다.

처음에는 “스마트시티 연구소에 코로나가 도착하기까지” 에 대해 시뮬레이션을 진행해보았습니다. ABM(Agent Based Modeling)이라는 모델링 방법을 사용하여 접근하였고, 미시적 행위자(사람들)의 상호작용으로 거시적 결과(전체 감염자 수, 병의 진행양상)가 어떻게 나타나는지 확인하기 좋은 방법이라고 생각했습니다. (ABM에 관해서는 기회가 되면 글을 또 올려보겠습니다.)

Netlogo라는 툴을 이용하여 시뮬레이션을 작성하였고, 연구소에 출근하는 직원들의 거주지 비율을 따져보면 대전이 가장 많았기에 대전, 세종, 그 외 지역으로 사람들의 속성을 나누었습니다. 감염률은 단순하게 “감염자 열 명을 만나면 한 번은 걸리겠지.” 라고 생각해서 0.1이라는 수치로 설정하였습니다.



정말 단순하게 설정하고 시뮬레이션을 돌렸지만, 불과 이틀만에 절반에 가까운 직원이 감염되는 결과를 보였습니다. 물론 실제로 이렇게 되지는 않았지만, 이 모형을 통해 **감염자 수가 늘어날 수록 전체 감염자 수가 증가하는 속도가 빨라진다**는 것을 알 수 있었습니다.

그러나, 모든 변수를 어림잡아 진행된 시뮬레이션이고 지역간 이동이나 질병이 진행되는 양상, 더 넒은 지역간 교통 등을 표현하기에는 Netlogo라는 학습용 언어로는 한계가 있었습니다.

## STEM을 이용한 실험

조금 더 나은 결과를 얻기 위해 어떻게하면 좋을지 고민해보던 중 연구소장님을 통해 STEM이라는 툴을 알게되었습니다. STEM은 IBM에서 처음 개발하고 지금은 Eclipse 재단에서 관리되고 있는 오픈소스 프레임워크 입니다. Spatio-Temporal Epidemiological Modeler의 약자로, 그대로 번역하자면 *시공간 역학 모델러* 정도 되는데 무슨 말인지 도무지 감이 오지 않습니다.

여기서의 역학(Epidemiology)은 전염병의 원인과 전염 과정에 대해 연구하는 학문을 의미합니다.

이해하기 쉽게 조금 풀어서 써보자면, 전염병의 전염 과정을 시간의 흐름과 공간 모형 위에 모형으로 표현할 수 있는 도구라고 정리할 수 있겠습니다.

### STEM의 주요 요소

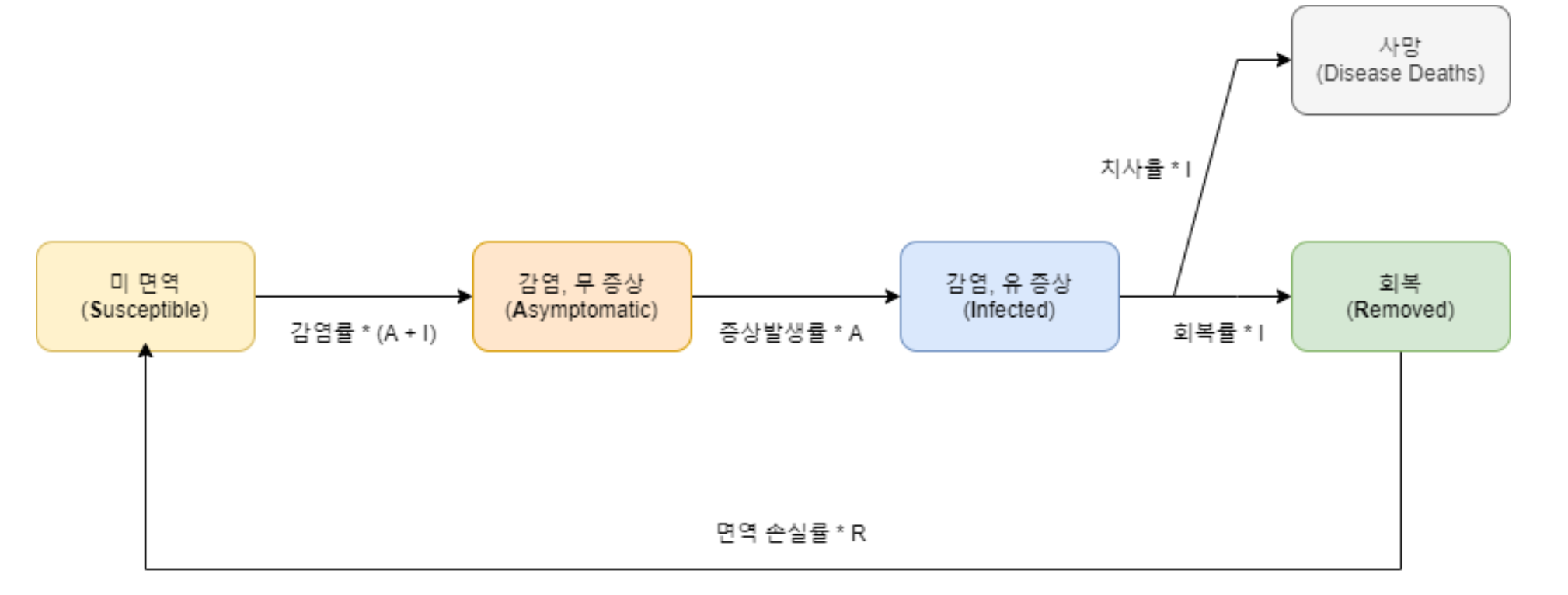
STEM에서 제가 사용했던 요소 중 중요한 것만 몇 가지를 소개하겠습니다. - Disease Model - 전염병의 진행 과정(상태)과 그것들을 이루는 속성을 표현하는 가장 중요한 모델 중 하나입니다. - STEM에서는 SIR, SEIR 모델 등 Multi Compartment Model을 기반으로 다양한 확장형 모델을 제공합니다. - Disease infector, initializer - 전염병의 시작을 표현하는 모델입니다. 위치(노드)와 초기 감염자 숫자(혹은 비율) 등을 지정할 수 있습니다 - Transportation Model - 인접 지역 간 인구의 이동을 표현하는 모델입니다. - 공통 경계를 가진 인접한 두 지역의 인구가 단위시간마다 얼마나 이동하는지를 나타냅니다. - Trigger, Predicator, Modifier - 특정 시간 혹은 특정 조건을 만족할 때(predicate) 모델의 변수를 수정할 수 있게(modify)합니다. - Trigger 안에 Predicator와 Modifier를 넣어 세 개의 요소를 한 쌍의 형태로 묶어 많이 사용합니다.

STEM은 국가의 지도와 지역 별 인구, 항공편 이동 등을 그래프 형식으로 표현해놓은 모델을 제공합니다. 때문에 위에 작성한 요소들만 적절히 이용하면 어렵지 않게 시뮬레이션을 구현할 수 있습니다.

### 코로나-19의 모형

잠깐, 코로나 얘기로 돌아와보겠습니다. 국내에서는 코로나로 인한 사망률이 다른 나라에 비해 많지 않은 편입니다. 높은 수준의 의료체계와 의료계 종사자 분들의 발빠른 대처가 더해진 결과이겠지요. 그럼에도 불구하고, 코로나-19의 무서운 점은 잠복기가 있고 잠복기 중에도 전염성을 갖는다는 것입니다. 그렇다면 이 코로나 바이러스는 어떻게 수학적으로 표현할 수 있을까요?

우선, 잠복기가 존재한다는 점은 감염가능군(Susceptible), 잠복기(Exposed), 감염군(Infected), 회복군(Recovered)로 표현되는 SEIR 모형과 비슷합니다. 그러나 기존의 SEIR 모델은 잠복기인 사람이 감염가능군에 있는 사람을 감염시킬 가능성이 없지만, 코로나는 잠복기인 사람도 전염성을 갖습니다. 또한, 회복이 되었던 사람도 재감염이 된 사례가 있습니다. 이러한 점들에 착안하여 아래와 같이 커스터마이징한 모델을 구현하였습니다.



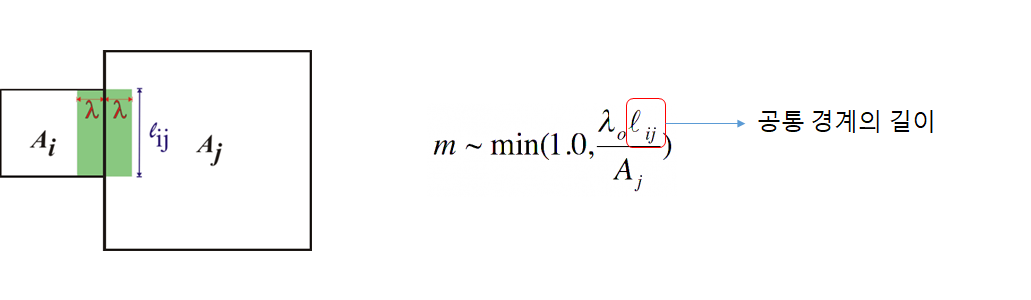
다른 상태(구획)로 변하는데에 영향을 미치는 주요 변수(parameter)들은 아래와 같습니다. - 감염률 - 증상발생률 - 회복률 - 치명률 - 면역 손실률

위의 변수들이 전염병의 각 구획 간의 관계를 표현하는 중요한 요소들입니다. 그러나, 공식적으로 발표된 자료가 없고 단순 통계자료만으로는 파악이 어렵기 때문에 우선 임의로 설정하고 다음 단계로 넘어가겠습니다.

### 대구, 경북 지역의 자발적 봉쇄

한국의 코로나-19 확산에서 꼭 언급되는 것은 바로 대구 신천지 집회입니다. 이 사건으로 인해 대량의 집단감염이 발생했고 대구, 경북지역의 많은 감염자를 낳게 되었습니다. 다른 지역으로 확산하는 것을 막고자 대구, 경북 시민들은 자발적 지역 봉쇄를 하기 시작했고 해당 지역 인구의 타 지역 이동이 최소화 되었습니다.

이것을 시뮬레이션에 반영하기 위해서는 우선 STEM에서 인구 이동을 어떻게 표현하는 지 알아야 합니다. 앞서, STEM의 주요 요소에서 언급한 Transportation Model이 인구 이동을 표현하는 모델입니다. 이 모형에서 인구 이동은 다음과 같이 표현됩니다.

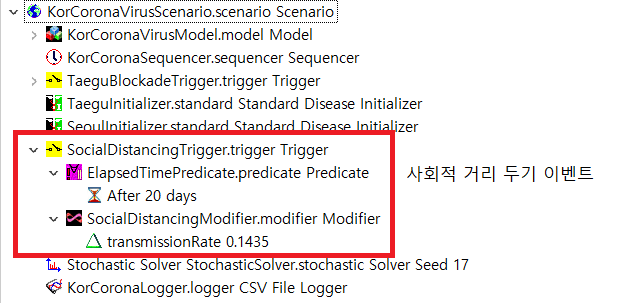


위 식을 통해 타 지역으로 이동하는 인구는 현재 있는 지역의 면적에 반비례하고, 기본 혼합거리(람다)와 공통 경계의 길이에 비례한다는 것을 알 수 있습니다. 이 식에서 공통 경계의 길이를 줄여서 대구, 경북의 자발적 봉쇄를 표현하였습니다.

### 사회적 거리 두기의 표현

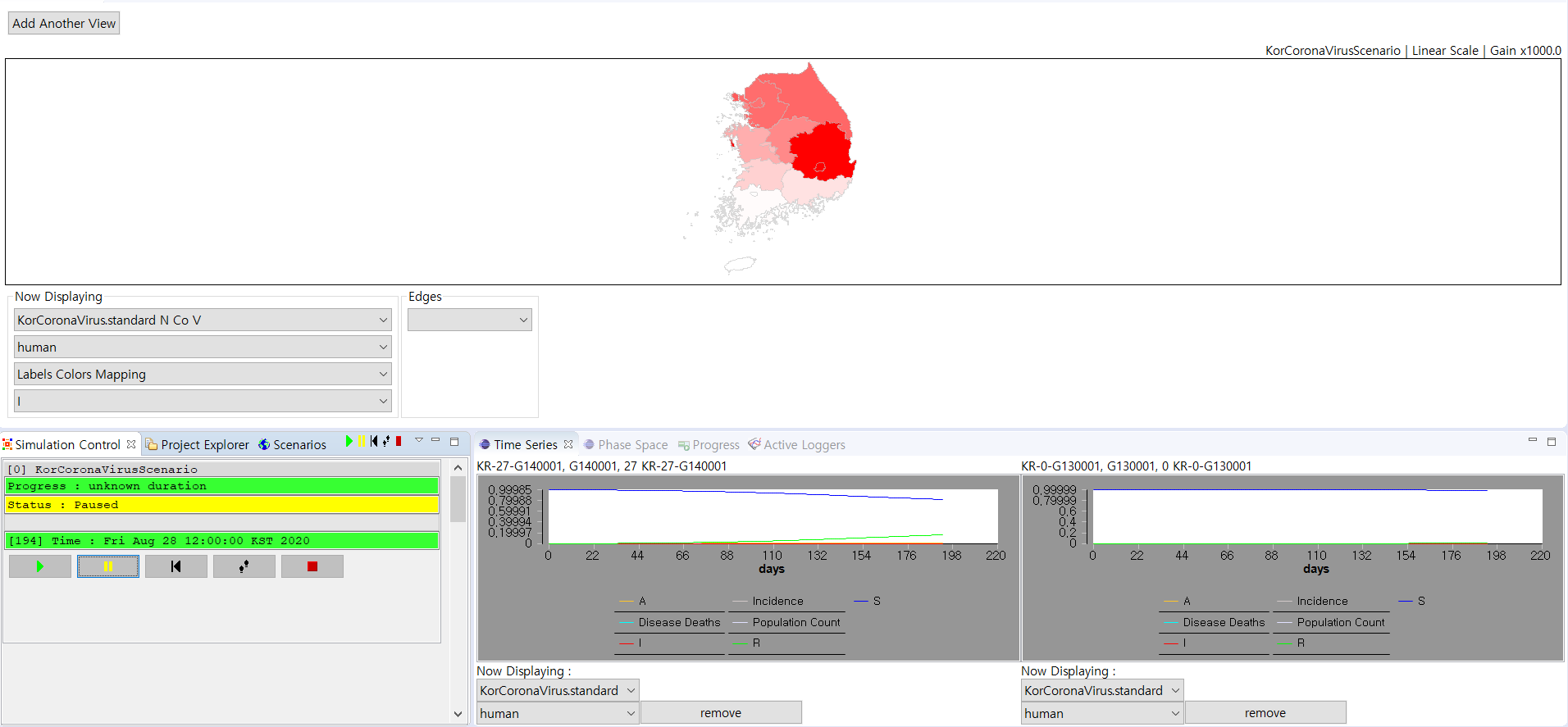
사회적 거리 두기는 감염 관리의 종류 중 하나이며 감염이 걸린 사람과 감염되지 않은 사람 사이의 접촉 가능성을 감소시켜 질병의 전파를 늦추는 것을 목표로 하는 방법입니다. 이러한 활동은 사람과 사람이 접촉할 확률 자체를 줄여주게 됩니다. 위에서 구현한 모형 중 감염률에 영향을 미치는 활동이죠.

우리 나라는 3월부터 본격적으로 강력한 사회적 거리 두기를 장려하며 코로나 확산을 늦추고 있습니다. 이를 STEM의 시뮬레이션에 반영하기 위해 3월부터 감염률을 직접적으로 낮추는 이벤트(Trigger, Predicate, Modifier)를 만들었습니다.



## 시뮬레이션 결과

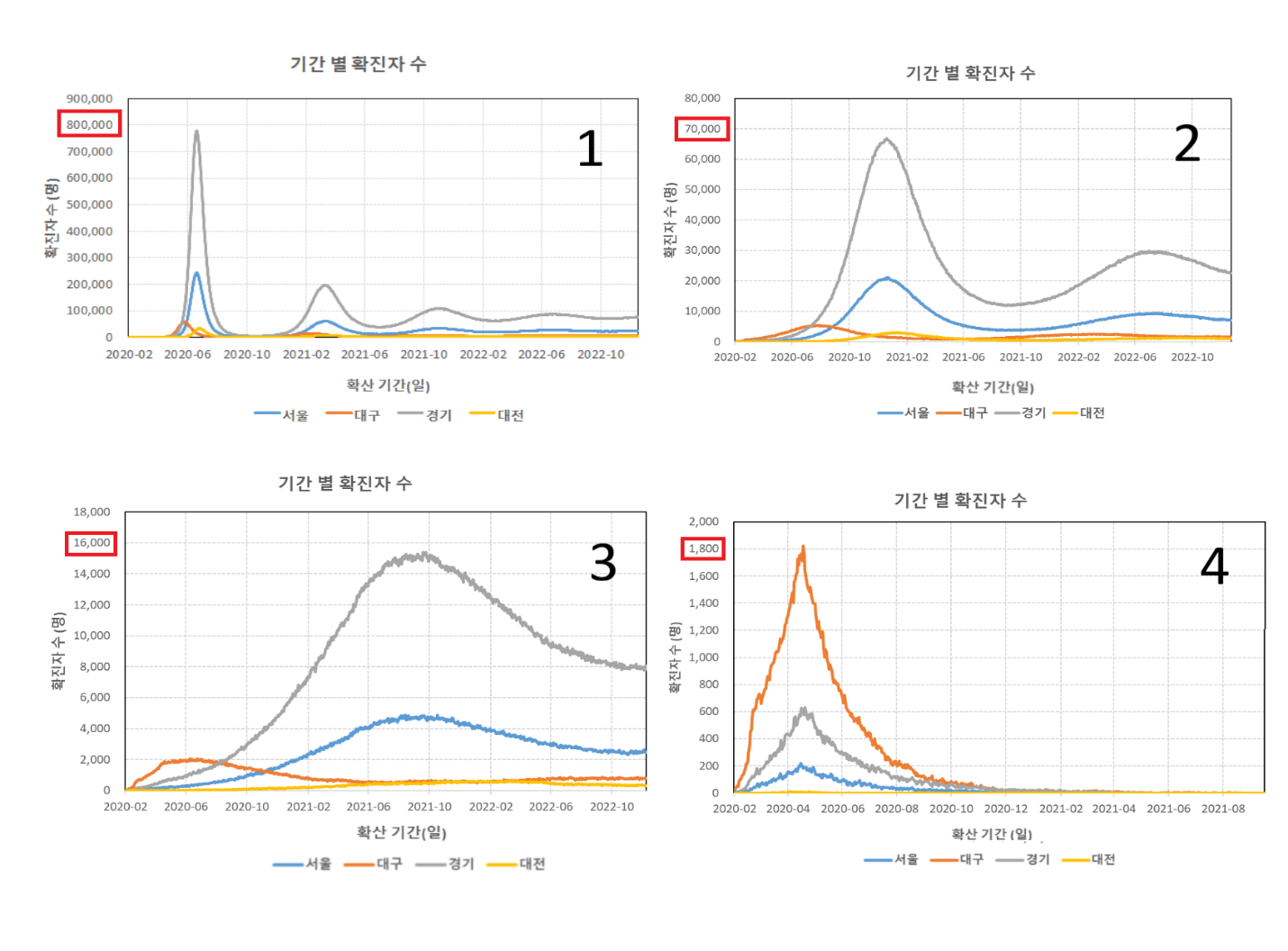
이번 시뮬레이션은 신천지 사건을 기점으로 작성했기 때문에 대구에 본격적으로 환자가 많아지기 시작하는 2월 17일을 기준으로 대구를 감염원으로 설정하고 시작했습니다. 시뮬레이션을 실행하게 되면 아래와 같은 화면을 보게됩니다.



기본적으로 STEM에서 제공하는 한국 지도에 감염자(I 구획)의 수가 많아질수록 해당 지역이 붉은 색에 가까워집니다. 아랫부분에 보이는 그래프는 지역 별 SAIR의 각 구획간 비율을 시간의 흐름대로 관측할 수 있는 그래프입니다.

여러 번 시뮬레이션을 실행하면서 4월 7일 기준 질병관리본부의 통계자료와 유사한 결과가 나올때까지 이벤트와 코로나 모델의 변수들을 조정하는 과정을 거쳤습니다. 지도로 보이는 결과를 통해 대구와 경북 지역의 자발적 봉쇄가 반영된 시뮬레이션임에도 불구하고 타 지역으로 감염이 시작되면 지역확산을 막을 수 없다는 것을 확인할 수 있었습니다.

약 1000일 정도 시뮬레이션을 진행했을 때 아래와 같은 결과를 얻을 수 있었습니다.



1번 : 사회적 거리 두기를 중간에 그만두는 경우

2번 : 사회적 거리 두기를 현재 강도로 지속하는 경우

3번 : 사회적 거리 두기를 조금 더 강화하는 경우

4번 : 사회적 거리 두기를 대폭 강화하는 경우

1번에서 3번으로 갈 수록 가장 높게 치솟는 구간의 수치가 낮아지고 옆으로 퍼지는 현상을 확인할 수 있습니다.

위 결과를 통해 아래와 같은 인사이트를 도출할 수 있었습니다.

1. (지도로 보이는 결과를 통해) 대구와 경북 지역의 자발적 봉쇄에도 불구하고 타 지역으로 감염이 시작되면 지역 확산을 막기 어렵다.

2. 사회적 거리 두기와 같은 활동들을 통해 확진자의 증가 시기를 늦추며 확진자의 수를 현저히 감소시킬 수 있다.

3. 일정 수치 이하로 감염률이 떨어지면 확산의 종식도 가능하다.

### 결과를 통해 얻은 개인적인 생각

지금도 강도 높은 사회적 거리 두기를 하고 있지만 여전히 확산이 되고있기 때문에 확산을 종식시키에는 많이 어렵다고 생각됩니다. 하지만 현재 상태를 지속하는 것만으로도 충분히 효과가 있고 한 시점에 확진자가 몰리지 않게 됨으로써 충분한 의료자원 확보와 백신 개발까지의 시간을 벌 수 있다고 생각합니다.

## 개선점

시뮬레이션을 통해 한 차례 결과를 얻고 보니 개선해야 할 점이 많이 느껴졌습니다.

첫 번째는 환자의 격리상태 표현입니다. 한국은 코로나 사태에 대한 대처가 세계적으로 가장 잘 되어있습니다. 신천지라는 집단 감염 사건이 있었음에도 불구하고 상대적으로 낮은 치사율과 높은 격리 해제율을 보여주고 있죠. 확진자의 격리 치료가 잘 되었기 때문이라고 볼 수 있는데요, 격리된 환자는 감염 모형에서 다른 구획으로 분리하여 미감염군이 감염될 확률에 영향을 미치지 못하도록 해야 하는데 시간 상의 한계로 구현하지 못한 점이 아쉽습니다.

두 번째는 교통 정보 반영입니다. STEM에서 기본으로 제공하는 지역 간 그래프와 인구 이동 모델은 우리 나라의 특성을 반영하기에는 부족한 점이 있습니다. 예를 들면 경기권에서 서울로 출퇴근 하는 인구가 굉장히 많기 때문에 수도권의 확진자 비율이 대구와 경북 다음으로 높아야 합니다. 그러나 STEM 시뮬레이션의 결과에서는 강원도와 비슷하게 확진자 비율이 나타나고 있습니다. 교통 정보를 반영한 한국의 모형을 만든다면 현실과 가까운 환경에서 시뮬레이션을 할 수 있다고 생각합니다.

세 번째는 인구 그룹의 특성 반영입니다. 현재 확진자를 연령대 별로 나누면 20대가 가장 높은 비율을 차지하고 있습니다. 그러나, 치사율은 70대 이상의 고령의 인구가 가장 높습니다. 또한 신천지와 같은 고 위험군 감염 그룹을 분리하여 변수를 별개로 설정하는 것도 중요합니다. 연령을 비롯해 치사율이나 감염률에 영향을 주는 그룹을 식별하여 반영한다면 정교한 시뮬레이션이 가능할 것이라고 생각합니다.

## 결론

스마트시티 연구소는 우리의 기술로 더 나은 세상이 되기를 바랍니다. 여러모로 보완할 점이 있지만, 더 안전하고 좋은 생활을 제공하기 위해 이런 도전적인 연구와 노력을 계속하겠습니다.

읽어주셔서 감사합니다 :)