

¿ES POSIBLE CONOCER LA EVOLUCIÓN DE UNA EPIDEMIA?

EL CASO DE COVID-19

A primera vista, el fenómeno de las epidemias parece ser de una enorme complejidad si tenemos en cuenta la cantidad de factores que pueden influir en su aparición y evolución. Sin embargo, en el proceso intelectual de identificar los mecanismos subyacentes y planificar estrategias de abordaje del problema podemos recurrir a un poderoso aliado: las matemáticas. Sus métodos, técnicas y procesos arrojan nuevas pistas para afrontar el estudio de un problema multifactorial.

En esta sección, *Preguntas en la Frontera*, se incluyen cuatro aportaciones internacionales de primer nivel que abordan esta cuestión desde distintos puntos de vista (sistemas complejos, fractales, inteligencia artificial...). En la primera, Marta C. Mora recurre al ámbito de los sistemas complejos y expone los fundamentos del modelo SIR, referencia ineludible en la temática. En la segunda, Anna y Robert Ziff, de Estados Unidos, comentan el análisis estadístico de los datos de expansión de la enfermedad y encuentran una estructura de crecimiento fractal.

La tercera aportación, realizada por un grupo de investigadores en China y Estados Unidos dirigidos por Zixin Hu, explica cómo aplicar técnicas de inteligencia artificial a la realización de predicciones en tiempo real de la evolución de los nuevos casos de la enfermedad. Por último, el grupo de Baoquan Chen, de la Universidad de Pekín, aborda el complejo tema de la visualización de datos públicos y el manejo de los mismos para facilitar el proceso de predicción.

ARTÍCULO LA MATEMÁTICA DEL CONTAGIO

MARTA C. MORA, UNIVERSITAT JAUME I (CASTELLÓN, ESPAÑA).

ARTÍCULO CINÉTICA FRACTAL DE LA PANDEMIA DE COVID-19

ANNA L. ZIFF Y ROBERT M. ZIFF, DUKE UNIVERSITY (DURHAM, EE.UU.) Y UNIVERSITY OF MICHIGAN (ANN ARBOR, EE.UU.)

ARTÍCULO PRONÓSTICOS PARA COVID-19 EN CHINA BASADOS EN INTELIGENCIA ARTIFICIAL

ZIXIN HU, QIFANG GE, SHIDI LI, LI JIN Y MINGMAO XIONG, FUDAN UNIVERSITY (SHANGHAI, CHINA) Y UNIVERSITY OF TEXAS (HOUSTON, EE.UU.)

ARTÍCULO ANÁLISIS VISUAL DE DATOS Y PREDICCIÓN DE LA SIMULACIÓN PARA COVID-19

BAOQUAN CHEN, MINGYI SHI, XINGYU NI, LIANGWANG RIAN, HONGDA JIANG, HEYUAN YAO, MENGDI WANG, ZHENGMIA SONG, QIANG BIAO Y TONG GE, UNIVERSIDAD DE PEKÍN (CHINA).

LA MATEMÁTICA DEL CONTAGIO

MARÍA C. MORA, UNIVERSITAT JAUME I (ESPAÑA)

En este artículo se exponen los fundamentos de la modelización matemática aplicada al estudio de pandemias, tomando como hilo conductor el modelo más sencillo y empleado, el modelo SIR, hoy expuesto como ejemplo en numerosos textos de ecuaciones diferenciales. A partir de este modelo, introduciremos algunas variaciones que nos permitan aplicarlo a casos reales y que aumenten su ámbito de aplicación.

INTRODUCCIÓN

El área que se ha venido a denominar los sistemas complejos es un área fascinante. Su objetivo último es el de ser capaz de representar modelos con muchos actores para los que las aproximaciones de estudio tradicionales son insuficientes. De un modo sorprendente, se constata que, con muy pocos parámetros, si un modelo de un fenómeno está bien planteado es capaz de reproducir gran parte de los comportamientos complejos que se observan en la realidad. Competición entre especies, extinciones, árboles filogenéticos, distribución de recursos naturales o fenómenos de percolación... el tipo de situaciones que pueden ser descritas a través del formalismo de estos sistemas es enorme y las conclusiones que arroja, fascinantes.

EL MODELO SIR

Para estudiar la propagación de una epidemia en términos de dinámica de sistemas, el modelo más sencillo que se puede emplear es el denominado modelo SIR, introducido hace casi un siglo por Kermack y McKendrick [1]. El modelo resulta muy simplificado y adolece de algunos problemas importantes. Probablemente el mayor de estos es que no concidea la posibilidad de

muerie de personas de una población, de modo que las epidemias nunca son fatales. Pero se pueden introducir modificaciones para tornarlo más realista. De momento, sepamos la población total de un entorno en tres categorías diferentes.

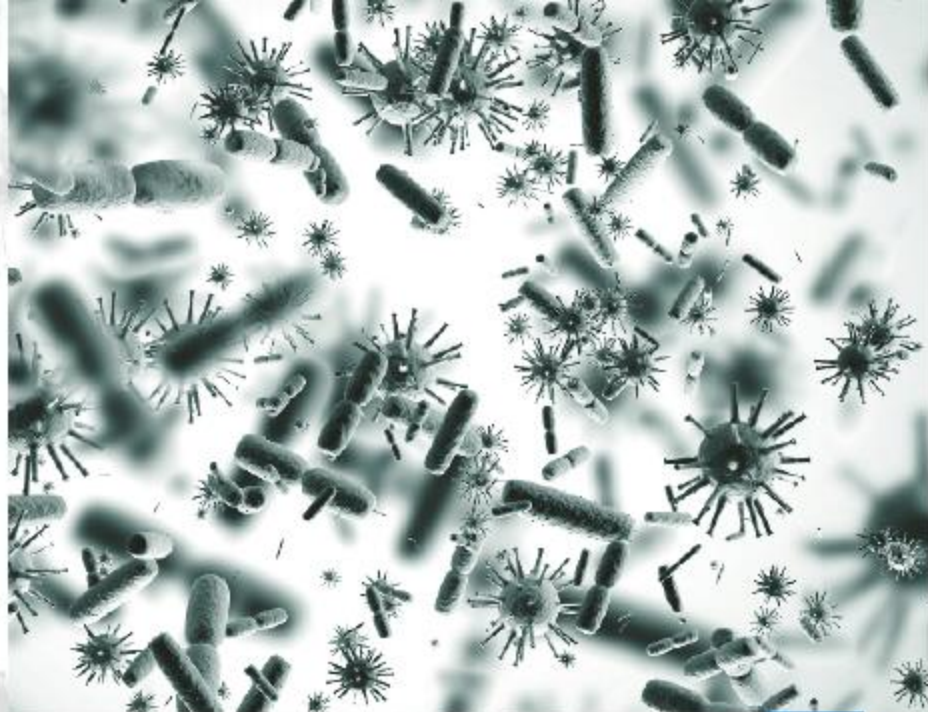
La primera categoría a considerar sería la de personas Sanas o *Susceptibles*, para las que podríamos emplear, en las fórmulas, la variable *S*. Estas personas todavía no han sido infectadas por el elemento patógeno. La segunda categoría sería la de personas *Infectadas*, a las que nos referiremos mediante la variable *I*. Por último, aquellas personas que ya han sufrido la enfermedad y que se han recuperado aparecen en la categoría de *Recuperadas*, identificadas mediante la variable *R*. El objetivo de la dinámica de sistemas consiste en obtener una serie de leyes que describan cómo varían las tres cantidades planteadas a lo largo del tiempo [2]. Veamos cómo es posible realizar esto.

Introduzcamos el parámetro β , que es la tasa de infección y modeliza el hecho de que una persona *Infectada* contagia a una *Sana*. Es sencillo llegar a la conclusión de que la tasa de personas *Infectadas* es mayor cuanto más gente *Sana* hay, cuanto más gente *Infectada* hay y, por supuesto, cuanto mayor probabilidad de infección. Eso puede expresarse por un simple producto de las tres cantidades:

$$\text{Tasa de Infección} = \beta \cdot S \cdot I \quad (1)$$

Por su parte, están los infectados que, eventualmente, se recuperan. La tasa de recuperación será mayor cuanto más infectados hay, y dependiente de un nuevo parámetro al que llamamos γ , que representa la probabilidad de que un infectado deje de estarlo.

$$\text{Tasa de Recuperación} = \gamma \cdot I \quad (2)$$



Si un modelo de un fenómeno está bien planteado es capaz de reproducir gran parte de los comportamientos complejos que se observan en la realidad.

Las cantidades introducidas afectan las tres poblaciones, hecho que se puede plasmar mediante una serie de ecuaciones diferenciales, que son ecuaciones que permiten establecer el cambio en función del tiempo. Para que el modelo sea asequible es necesario asumir que las poblaciones solo dependen de estos parámetros, y no de otros externos. Por ejemplo, no se tiene en cuenta la existencia de nacimientos o fallecimientos en el sistema. Estas suposiciones, que a primera vista pueden parecer muy restrictivas, no lo son tanto si se consideran que son fenómenos que se desarrollan en distintas escalas de tiempo. En general, la propagación de una enfermedad es un fenómeno relativamente rápido comparado con las muertes y los nacimientos "naturales".

Así pues, la población de personas *Sanas* decrece con la tasa de infección. La población de personas *Infectadas* varía por dos motivos: a) crece con la tasa de infección, b) decrece con la tasa de recuperación. Por último, la población de personas *Recuperadas* crece con la tasa de recuperación.

LO QUE DICEN LAS ECUACIONES

Todos los razonamientos anteriores llevan a un sistema de tres ecuaciones diferenciales que constituye una descripción elegante de un fenómeno tan complejo como la expansión de una epidemia, pero, ¿qué tiene que decirnos? Para responder a esta pregunta es necesario resolver el sistema de ecuaciones, asunto que no siempre es sencillo. Las ecuaciones obtenidas son no lineales, lo que en extremo el trabajo analítico, dado que no pueden resolverse de manera directa.

Por fortuna, desde hace mucho tiempo, los matemáticos han conseguido lidiar con este tipo de problemas realizando aproximaciones matemáticas, las cuales consisten, básicamente, en dividir el tiempo y el espacio en intervalos diminutos y aplicar determinadas aproximaciones a las ecuaciones para que sean sencillas de tratar [3]. El proceso se muestra especialmente beneficioso cuando, además, se recurre a máquinas de cómputo de elevadas prestaciones, como los ordenadores, capaces de realizar multitud de cálculos en un breve lapso temporal.

PARA ESTUDIAR LA PROPAGACIÓN DE UNA EPIDEMIA EN TÉRMINOS DE DINÁMICA DE SISTEMAS, EL MODELO MÁS SENCILLO QUE SE PUEDE EMPLEAR ES EL DENOMINADO MODELO SIR, INTRODUCIDO HACE CASI UN SIGLO POR KERMACK Y MCKENDRICK



El fenómeno de las pandemias presenta una enorme complejidad si tenemos en cuenta la cantidad de factores que influyen en ella.

El campo del cálculo numérico presenta una ferviente y muy destacada actividad en la actualidad [4].

La resolución de estas ecuaciones deja claro que el comportamiento cuantitativo del sistema depende en gran medida de los valores asignados a los parámetros β y γ . Pero cualitativamente, el comportamiento es fácilmente interpretable y se puede ver en la Figura 1. En ella se ve cómo la población de infectados crece hasta un máximo, momento a partir del cual decrece hasta desaparecer. En el estado de equilibrio, la población total queda reducida a dos categorías: aquellos que sufrieron la enfermedad y los que no.

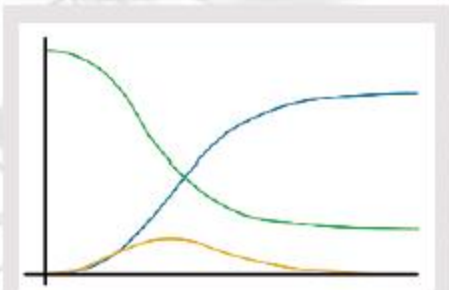


Figura 1. Comportamiento típico del modelo SIR. La traza verde corresponde a los individuos sanos. La traza azul representa a la población recuperada. La traza naranja describe a los individuos infectados.

LAS APROXIMACIONES NUMÉRICAS CONSISTEN, BÁSICAMENTE, EN DIVIDIR EL TIEMPO Y EL ESPACIO EN INTERVALOS DIMINUTOS Y APLICAR DETERMINADAS APROXIMACIONES A LAS ECUACIONES PARA QUE SEAN SENCILLAS DE TRATAR

CONCLUSIÓN

El sencillo modelo SIR puede ser completado para introducir la probabilidad de muerte de los infectados [5]. En cualquier caso, el ejemplo expuesto ya deja patente que, con muy pocos parámetros, es posible captar la esencia de un fenómeno masivo, tristemente presente en el mundo actual y de consecuencias a menudo devastadoras. En nuestra mano está invertir esfuerzos en mejorar el grado de conocimiento del fenómeno a través de su estudio y modelización matemática.

REFERENCIAS

- [1] W. G. Kermack, A. G. McKendrick, *Proceedings of the Royal Society of London Series A*, 115 (1927) 700-721.
- [2] J. Arino, G. Gendrik, *Dinámica de sistemas*, Alianza, Madrid, 1998.
- [3] D. Hughes-Hallett et al, *Applied Calculus*, 2nd ed., Wiley, Toronto, 2002.
- [4] C. F. Gerald, P. O. Wheatley, *Applied numerical analysis*, 7th ed., Pearson, Londres, 2003.
- [5] J. G. Sogarra, *Vida artificial del caos al orden*, Algor, A Lima, 2002.

LA AUTORA

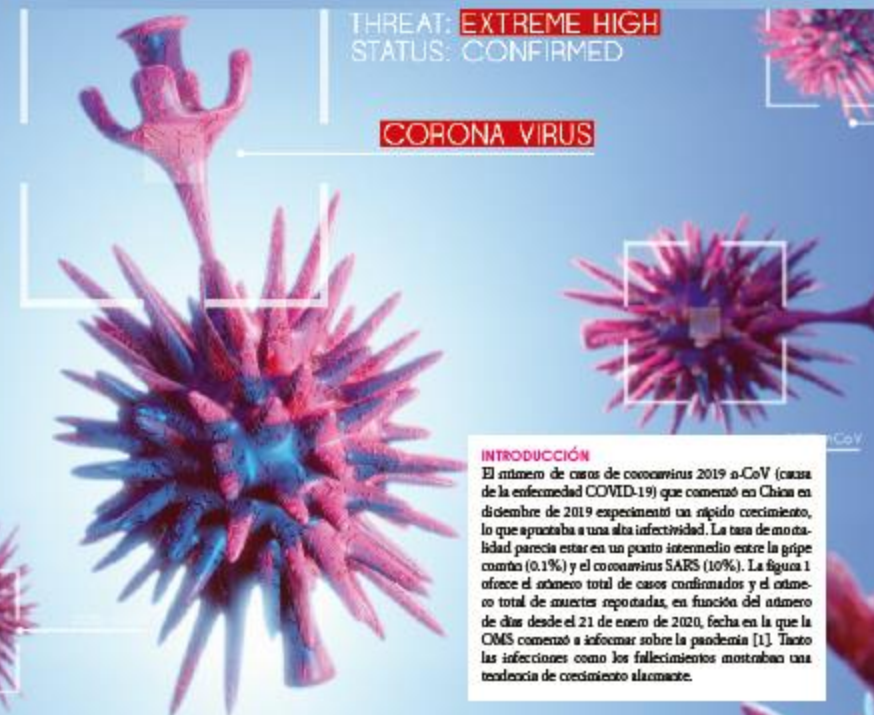
Marta Covadonga Mora Aguilar

Doctora Ingeniera Industrial por la Universitat Politècnica de València. Desde 2006 es profesora del Departamento de Ingeniería Mecánica y Construcción de la Universitat Jaume I, en Castellón, donde compatibiliza sus obligaciones docentes e investigadoras. En el término de la investigación ha realizado aportaciones en los campos de la robótica, la automatización industrial y, más recientemente, en biomatemática.

CINÉTICA FRACTAL DE LA PANDEMIA COVID-19

ANNA L. ZIPP Y ROBERT AL. ZIPP, DUKE UNIVERSITY Y UNIVERSITY OF MICHIGAN (EEUU)

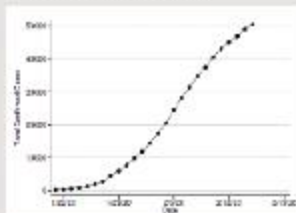
El enfoque epidemiológico clásico en el estudio de una pandemia como la de COVID-19 consiste en cuantificar un número de reproducción y tiempo de infección. Sin embargo, esta suposición conduce a un crecimiento exponencial y, a pesar de que en el caso indicado la tasa de crecimiento es alta, se constata que no sigue dicho patrón. En este artículo se observa el crecimiento como una ley de potencias, indicativa de un fractal subyacente o una red de conexiones entre individuos susceptibles e infectados.



INTRODUCCIÓN

El número de casos de coronavirus 2019 n-CoV (causa de la enfermedad COVID-19) que comenzó en China en diciembre de 2019 experimentó un rápido crecimiento, lo que apuntaba a una alta infectividad. La tasa de mortalidad parecía estar en un punto intermedio entre la gripe común (0.1%) y el coronavirus SARS (10%). La figura 1 ofrece el número total de casos confirmados y el número total de muertes reportadas, en función del número de días desde el 21 de enero de 2020, fecha en la que la OMS comenzó a enfocarse sobre la pandemia [1]. Tanto las infecciones como los fallecimientos mostraban una tendencia de crecimiento exponencial.

(a) Número de casos confirmados



(b) Número de muertes

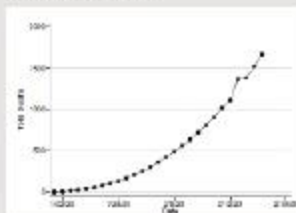


Figura 1. Infecciones y muertes por COVID-19 desde el 22 de enero de 2020. Datos de la OMS

casos [2]. Esto supone que no hay inhibición debido a la interacción con personas ya infectadas, cuarentenas u otras medidas profilácticas. Si bien los datos muestran un gran crecimiento, también queda patente que no siguen un comportamiento exponencial. De hecho, los datos casan correctamente si se asume un modelo de ley de potencias de exponente mayor que dos, como se muestra en la figura 2. El valor no entero de esta ley sugeriría un tipo de comportamiento fractal del mecanismo de susceptibilidad.

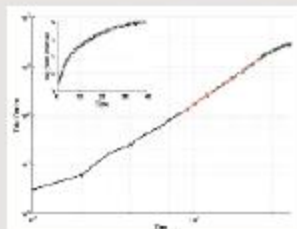


Figura 2. Muertes totales en función del tiempo desde el 21 de enero de 2020. La línea roja en el diagrama grande corresponde a una ley de potencias. Datos de la OMS.

PRIMER ANÁLISIS EPIDEMIOLÓGICO

En el análisis epidemiológico estándar se parte del supuesto de que el número de casos en este tipo de epidemias aumenta exponencialmente, en base a una tasa fija de reproducción. Varios estudios sobre COVID-19 han adoptado este enfoque para modelar el número de

Hay que hacer notar que el número total de muertes en la figura 1 parece estar creciendo en forma aproximadamente cuadrática, mientras que el número de casos confirmados que se muestra en la misma figura parece seguir un comportamiento similar, pero se está desacelerando en los últimos datos. Podría tratarse de un error de medición debido a múltiples factores, que incluyan los fallos de representación del número de casos confirmados y la variación de los criterios utilizados para identificación de la enfermedad. No es el objetivo de nuestro estudio el de cuantificar cualquier sesgo intencional en los datos informados, pero sí es interesante destacar que, a partir del 7 de febrero del 2020, los pacientes que dan positivo en COVID-19 pero no muestran síntomas de la enfermedad, ya no se contabilizan como casos confirmados [3].

Así pues, se centró el análisis en el número de muertes para paliar el efecto de las ambigüedades en la contabilización de casos confirmados, suponiendo que las muertes son una indicación más confiable de la extensión de la enfermedad. Debemos recalcar que, aunque mejor, tampoco se trata de un indicativo perfecto, ya que algunas muertes pueden ser achacadas a otras causas, tales



La aparición de un exponente fraccional y el comportamiento de la ley de potencias sugiere un proceso fractal subyacente.



como neumonía severa, y no atribuidas a este virus, por lo que estos números también pueden ser inexactos. Aun así, el conjunto de datos debería dar una buena indicación de la tendencia general de la pandemia. Se encuentra que una ley de potencias proporciona un mejor ajuste a los datos que una simple función cuadrática, tal como ya fue postulado por Brandeburgo [4].

PROCESO FRACTAL

La aparición de un exponente fraccional y el comportamiento de la ley de potencias sugiere un proceso fractal subyacente, con base en una especie de red de interacción en un mundo cada vez más pequeño, donde las personas tienen muchos vecinos locales y conexiones ocasionales de largo alcance (como las causadas por personas que viajan en trenes, barcos y aviones). Los resultados implican que esta red tiene una dimensión fractal de trayectoria mínima efectiva de 2,25. Esto implica que la enfermedad produce picos en los que el crecimiento es momentáneamente exponencial, pero luego se ralentiza hasta otro pico. El efecto promedio de este crecimiento apocadamente produce una ley de potencias.

Hacia mediados de febrero se observó una caída en el crecimiento de nuevos casos confirmados de China, lo que sugiere una atenuación de la pandemia. Sin embargo, dicha atenuación no se reflejaba en la tasa de mortalidad. Cabeen dos posibles explicaciones para esta discrepancia: (1) el número de casos confirmados puede estar mostrando la variabilidad tal como ha sido en el pasado, y los números actuales no capturan la situación real, o (2) la velocidad de la pandemia está, de hecho, disminuyendo, pero llevará un tiempo (del orden de una semana) para que dicha tendencia se muestre en los datos de fatalidad. Por supuesto, la segunda posibilidad sería la más deseable, pero no se puede dar por hecho, habida cuenta de los problemas en los informes. Las desviaciones por encima de este

comportamiento de ley de potencias podría indicar que la pandemia se está expandiendo desde el actual nivel de control, mientras que las desviaciones por debajo pueden indicar que la enfermedad está empezando a desaparecer.

EVOLUCIÓN DE LA ENFERMEDAD

A partir de marzo del 2020 el número de muertes en China experimentó una rápida disminución, lo que se adaptaría bastante bien a un decrecimiento exponencial de la ley de potencias similar al predicho por el modelo de red de Vázquez [5]. Al mismo tiempo, vemos que las muertes fuera de China aumentan rápidamente, y de un modo similar al comportamiento temporal de las estadísticas de China. Por lo tanto, vemos tres etapas de la propagación de la enfermedad en términos de número de muertes: crecimiento exponencial, comportamiento como una ley de potencias y posterior disminución exponencial.

No es descartable que, a raíz de los viajes internacionales normales que se producen en nuestra sociedad, el virus se vaya extendiendo a muchas partes del mundo y por lo que surgen focos. Ello nos lleva a un escenario de pandemia global. □

REFERENCIAS

- [1] World Health Organization, Technical Report 124, WHO <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>, 2020.
- [2] S. Zhao et al., *International Journal of Infectious Diseases*, 92 (2020) 214-217.
- [3] *Cochin Science Magazine*, February 7th, (2020).
- [4] A. Brandeburgo, *arXiv:2002.03638 [q-bio]*.
- [5] A. Vázquez, *Phys. Rev. Lett.*, 96(3) (2006) 038702.

LOS AUTORES

Anna L. Ziff
 Es Ph.D. Cand. de Economía y asistente de investigación en el Department of Economics de la Duke University en Durham (EE.UU.). Es graduada en Estadística y Políticas Públicas por la University of Chicago donde trabajó como Research Professional.

Robert M. Ziff
 Es Master en Ciencias Físicas por la University of California en Los Angeles y Doctor en Física Estadística por la Rockefeller University. Actualmente es Profesor en el Center for the Study of Complex Systems y el Department of Chemical Engineering de la University of Michigan E.E.U.U.

PRONÓSTICOS PARA COVID-19 EN CHINA BASADOS EN INTELIGENCIA ARTIFICIAL

POB ZIXIN HU, QIYANG GE, SHUOLI LI, JIN, Y MOMIAO XIONG,
FUJIAN UNIVERSITY (SHANGHAI, CHINA) Y UNIVERSITY OF TEXAS (HOUSTON, EEUU)

Como alternativa a los modelos epidemiológicos de dinámica de transmisión de COVID-19 en China, se propone el uso de métodos basados en inteligencia artificial para la realización de predicciones en tiempo real sobre la extensión y duración de la pandemia en China. Para ello utiliza un codificador automático apilado y modificado para modelar las dinámicas de transmisión y se aplica a los datos ofrecidos por la OMS, lo que conduce a predicciones con errores menores del 3%.

INTRODUCCIÓN

En diciembre de 2019, cerca de la ciudad de Wuhan, el virus COVID-19 dio un salto de una especie animal al ser humano, iniciando un brote de enfermedad respiratoria. Aunque los modelos epidemiológicos son útiles para estimar la dinámica de transmisión, dirigir los recursos y evaluar el impacto de las estrategias de intervención, requieren parámetros y dependen de muchos supuestos, lo que hace que las predicciones se desvíen mucho de los datos reales [1-3]. Para superar las limitaciones del enfoque basado en modelos epidemiológicos y contribuir a la planificación de estrategias y políticas de salud pública, los autores desarrollaron un método basado en inteligencia artificial (IA) para prever, en tiempo real, nuevos casos de COVID-19 [4]. El análisis se basa en los datos de vigilancia de los casos confirmados de COVID-19 en China hasta el 8 de febrero de 2020.

FUENTE Y ANÁLISIS DE DATOS

Los datos sobre casos confirmados de COVID-19 se obtuvieron de los informes publicados por Suring News Network (<https://www.thepaper.cn/>) y la OMS (<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>). Estos datos resumían el número total acumulado y los nuevos casos en China distribuidos en 31 provincias o ciudades dentro de la parte continental de China y tres otras regiones (Hong Kong, Macao y Taiwán). Los datos fueron organizados en forma de matriz. Cada fila, correspondiente a una región geográfica, suponía, pues, una serie temporal con los nuevos casos confirmados cada día. De este modo, la matriz tenía un tamaño de 34 filas por n columnas, donde n es el número de días tomados en consideración.

Para el modelado de la serie temporal se desarrolló un codificador automático modificado (MAE, por sus siglas en inglés, *Modified Auto-encoder*) [5][6]. A diferencia del codificador automático clásico, donde el número de nodos en las capas generalmente disminuye de la capa de entrada a las capas latentes, los números de los nodos en la entrada, la primera capa latente, la segunda capa latente y las capas de salida en el MAE fueron 8, 32, 4 y 1, respectivamente, tal como puede verse en la figura 1.

PARA EL MODELAO DE LA SERIE TEMPORAL SE DESARROLLÓ UN CODIFICADOR AUTOMÁTICO MODIFICADO (MAE, POR SUS SIGLAS EN INGLÉS, MODIFIED AUTO-ENCODER).

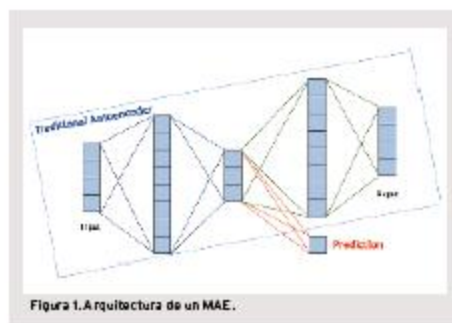


Figura 1. Arquitectura de un MAE.

Se define una muestra como una serie formada por 8 valores diarios, y se consideraron 128 muestras como conjunto de entrenamiento. Para ello se parte de la matriz de datos, seleccionando aleatoriamente el día de inicio del segmento y tomando los 7 sucesivos. Los pesos del MAE se estimaron utilizando un algoritmo básico de retropropagación, y se repitió el proceso de entrenamiento 5 veces. La predicción final se comparaba con los datos confirmados. Además, se empleó un pronóstico recursivo en múltiples pasos, consistente en usar un modelo de un paso varias veces. De esta manera, la predicción obtenida en un paso se utiliza como entrada para la predicción del paso siguiente.

RESULTADOS

La figura 2 muestra las curvas de número total de los casos confirmados diarios y acumulados de COVID-19 en China en función del tiempo. Los casos reportados van

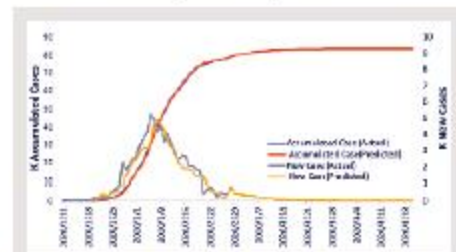


Figura 2. Curvas de nuevos casos diarios y casos acumulados de COVID-19 en China en función del tiempo, del 11 de enero de 2020 a 20 de abril de 2020. Se ofrecen tanto los valores medidos como los predichos por el MAE.

del 11 de enero de 2020 hasta el 27 de febrero del mismo año. Tras el proceso de estimamiento, el predictor empezó a funcionar el 20 de febrero de 2020. Como puede verse en la figura, la curva de pronóstico está muy cerca de la obtenida con los datos confirmados. De hecho, se aprecia que el pico de nuevos casos confirmados tiene lugar el 5 de febrero. Desde el momento que el predictor entró en funcionamiento, la magnitud del error se mantuvo por debajo del 3%.

La figura 3 particulariza las predicciones para cada una de las regiones en las que se dividió el estudio. Se observa que el momento en que la evolución de la enfermedad entra en un estado de meseta es diferente para cada región, como parece lógico pensar. La primera región que entró en zona de meseta fue Xizang, en 21 de enero, seguida de Macao (26 de enero) y Qinghai (28 de ene-

ro). El predictor anticipa que una de las últimas regiones en entrar en la región será Hubei, y que este comportamiento no llegará hasta mediados de abril. También se pueden observar las diferencias de las curvas entre las distintas regiones, que implican diferentes patrones dinámicos de la transmisión del virus.

CONCLUSIÓN

Como alternativa al modelo de transmisión epidemiológica, utilizamos MAE para pronosticar la trayectoria en tiempo real de la dinámica de transmisión y generar predicciones de COVID-19 en las regiones de China. La precisión de los métodos basados en IA es destacable. Si los datos son fiables y no hay segunda transmisión, es posible predecir con exactitud la dinámica de transmisión de un virus en las distintas regiones, lo que aporta indicios para ayudar a la planificación de políticas de salud pública.

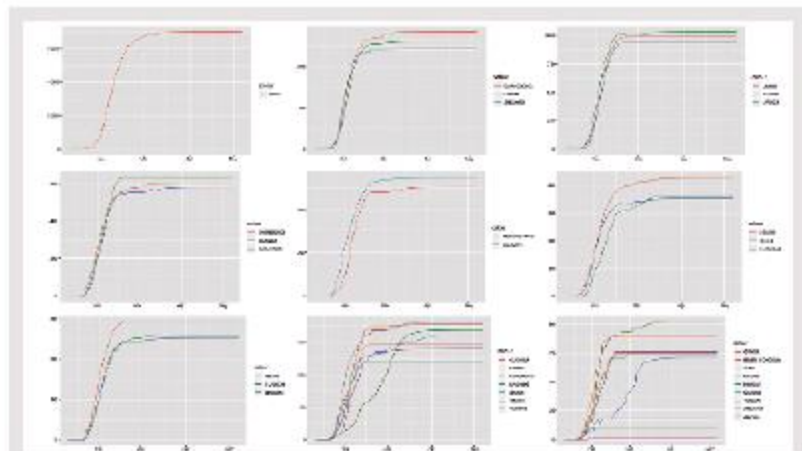


Figura 3. Curvas de predicciones de casos acumulados de COVID-19 en las distintas regiones de China en función del tiempo, del 11 de enero de 2020 a 20 de abril de 2020.

REFERENCIAS

[1] Li Q, et al, *NEngl J Med*, 2020, 29 enero (2020)
 [2] Wu JT, Leung K, Leung GM, *Lancet*, 29 enero (2020) 388
 [3] Zhang, Musa, Lin, et al, *J Clin Med*, 9(2) (2020) 388
 [4] S. Funk et al, *Epidemics*, 22 (2018) 56-61
 [5] D. Charle et al, *Information Fusion*, 44 (2018) 78-96
 [6] Yuan X et al, *IEEE Trans on Industrial Informatics*, 14(7) (2018) 3235-3243

LOS AUTORES

Zhin Hu y U Jin trabajan en el State Key Laboratory of Genetic Engineering and Innovation Center of Genetics and Development de la School of Life Sciences en la Fudan University de Shanghai, China, y también colaboran con el Human Phenome Institute de la misma Universidad. Qiyang Ge es miembro de The School of Mathematic Sciences en la Fudan University de Shanghai, China. Shuili Li y Momo Xiong son miembros del Department of Biostatistics and Data Science de la School of Public Health de The University of Texas Health Science Center at Houston, Houston, EEUU.

ANÁLISIS VISUAL DE DATOS Y PREDICCIÓN DE LA SIMULACIÓN PARA COVID-19

BACQUAN CHEN, MINYI SHI, XINGYU NI, LIANGWANG RUAN, HONGDA JIANG, HEYUAN YAO, MENGDI WANG, ZHENGHUA SONG, QIANG ZHOU Y TONG GE, UNIVERSIDAD DE PEKÍN, CHINA.



La epidemia de COVID-19 se ha convertido en una pandemia global según la OMS. China ha sido el país más afectado desde el brote del virus, que algunos expertos remontan a finales de noviembre. El 23 de enero, el gobierno de Wuhan reconoció la gravedad de la epidemia y tomó la drástica medida de cerrar todo transporte con el exterior. En este artículo se intenta responder a varias cuestiones clave: ¿cómo se propagó el virus desde el epicentro de la ciudad de Wuhan al resto del país? ¿En qué medida las resoluciones tomadas ayudaron a controlar la situación? Y más importante aún, ¿es posible pronosticar el desarrollo futuro del evento bajo ciertos cambios de condiciones?

INTRODUCCIÓN

La epidemia de COVID-19 se extendió desde Wuhan a sus áreas circundantes, y posteriormente al resto del país, especialmente las grandes áreas metropolitanas, como Beijing, Shanghai y Guangzhou. El patrón de transmisión refleja el gran movimiento de población previa al Año Nuevo chino (25 de enero). El 23 de enero de 2020, Wuhan decretó una medida extrema de cuarentena total, suspendiendo todos los transportes que entran o salen de la ciudad. Es posible, partir de varias fuentes públicas, visualizar los datos de propagación de la infección, tal como se ofrecen en la figura 1. Hay una gran correlación entre los movimientos migratorios y la tasa de infección, sobre todo en las primeras etapas de la enfermedad.

EL MODELO SEIR NO CONSIDERA LAS MEDIDAS DE CONTROL TALES COMO LA CUARENTENA, LO QUE NOS LLEVÓ A PROPONER UN NUEVO MODELO DE DINÁMICA EPIDÉMICA, EL C-SEIR



Figura 1. (Izquierda) Flujo de población desde Wuhan al resto de China. (Centro) Casos diagnosticados de infección a 31 de enero de 2020. (Derecha) Casos diagnosticados de infección a 9 de febrero de 2020.

EL MODELO C-SEIR

El 31 de enero, *The Lancet* publicó un artículo [1] sobre predicción y control de la propagación de la enfermedad COVID-19 en China mediante el modelo clásico SEIR, que divide a la población en cuatro categorías: susceptibles (S), expuestos (E), infectados (I) y recuperados (R). El modelo asume que los individuos pasan de una categoría a otra con una cierta probabilidad. A partir de este modelo, los autores estimaron que el número de personas infectadas en Wuhan habría alcanzado alrededor de 75.815 el 25 de enero. El artículo básicamente describía el efecto de la cuarentena de Wuhan en toda la ciudad y otras medidas de control agresivas en todo el país. El pronóstico presentaba un escenario grave tanto en China como a nivel mundial, lo que fue útil para alertar al público, pero los datos ofrecidos estaban bastante alejados de la realidad incluso en el momento de su publicación.

El modelo SEIR no considera las medidas de control realizadas, tales como la cuarentena, lo que nos llevó a proponer un nuevo modelo de dinámica epidémica, el C-SEIR [2], el cual añade dos nuevos grupos en cuarentena en SEIR: el grupo en cuarentena por sospecha (P), y el grupo en cuarentena por diagnóstico (Q). Los individuos

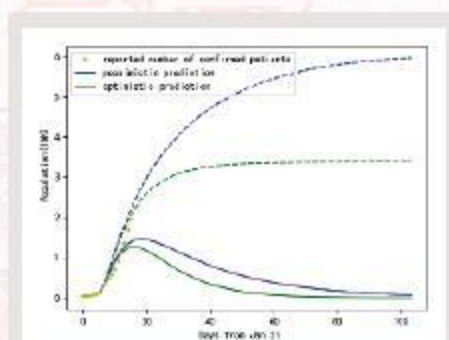


Figura 2. Predicción de infecciones en la provincia de Hubei. Las curvas verde y azul muestran resultados de dos conjuntos de parámetros, que representan escenarios optimistas y pesimistas respectivamente.

en los grupos P y Q no tienen la capacidad de transmitir el virus. A partir de los datos recolectados en la provincia de Hubei, de la que Wuhan es capital, fue posible calcular los parámetros del modelo y ejecutar una simulación. Las predicciones obtenidas se resumen en la figura 2.

De las simulaciones se desprende la importancia de una cuarentena adecuada en la etapa inicial de la epidemia. Como puede verse a partir de la figura 3, al cambiar la fecha de cuarentena en solo 2 días se introduce un factor 2 en el número de personas infectadas, tanto en sentido positivo como negativo. Un número demasiado grande de personas infectadas puede llegar a colapsar el sistema médico, situación que se dio en Wuhan.

CONCLUSIONES

La simulación basada en el modelo C-SEIR ha demostrado una efectividad significativa de las medidas de cuarentena adoptadas en China. Si bien se reconoce el éxito de estos esfuerzos, hay que señalar que el coste de las mismas quizá ha sido demasiado elevado. En retrospectiva, habría sido posible obtener resultados similares mediante una cuarentena más adecuada (menos estricta) con un inicio anterior. Dada la expansión de COVID-19 a otras áreas geográficas, y tomando como base nuestras observaciones y subsecuente análisis, podemos realizar las siguientes sugerencias:

- 1) Alertar y educar al público sobre el virus y su prevención.
- 2) Monitorizar de cerca el desarrollo de la epidemia diagnosticando rápidamente a las personas infectadas con síntomas característicos, ponerlas en cuarentena y tratarlas.
- 3) Alertar a quienes hayan tenido contacto cercano con las personas infectadas y asegurarse limitar su actividad social, incluso si no presentan síntomas, que realicen una cuarentena si presentan síntomas leves, comuniquen con las autoridades sanitarias para recibir un diagnóstico si presentan síntomas característicos. Los datos indican que el tiempo adecuado de cuarentena es de dos semanas.

Esta tercera medida refleja la estrategia de cuarentena del modelo C-SEIR, y ha demostrado su eficacia tanto en la simulación del modelo como en la verificación de la realidad.

Recientemente se ha reportado una tasa de mortalidad del 1,4% para el virus [3]. Pero fuera de la provincia de Hubei, la tasa de mortalidad parece ser mucho más baja, llegando incluso al 0,85%. Se puede especular con que la letalidad del virus disminuye con una mayor transmisión. Pero los datos indican que COVID-19 es definitivamente más severa que una gripe por lo tanto, requiere disponer de precauciones adicionales y recursos médicos adecuados.

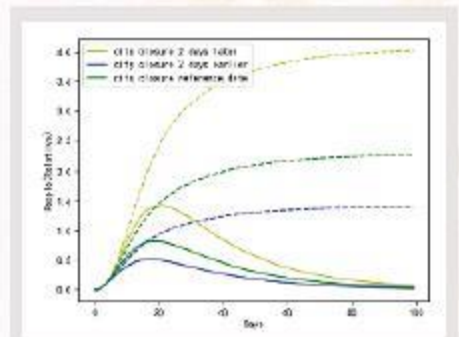


Figura 3. La simulación sugiere que el ejercicio de la cuarentena en la etapa temprana del brote del virus puede conducir a una disminución o aumento drástico de las infecciones totales.

DE LAS SIMULACIONES SE DESPRENDE LA IMPORTANCIA DE UNA CUARENTENA ADECUADA EN LA ETAPA INICIAL DE LA EPIDEMIA

LOS AUTORES

Baoquan Chen, Mingyi Shi, Xinyu Ni, Lijiangwan Ruan, Hongda Jiang, Haiyan Yao, Hongdi Wang, Zhonghua Song, Qiang Zhou y Tong Ge pertenecen al Centro en Fronteras de Computing Studies, de la Universidad de Pekín, China. El director del grupo, el profesor Chen, es Doctor en Ciencias Informáticas por la State University of New York (EE.UU.). Actualmente es Director Ejecutivo del Centro en Fronteras de Computing Studies de la Peking University, habiendo sido con anterioridad, entre otros prestigiosos cargos, Decano de la School of Computer Science and Technology de la Shandong University (China).

REFERENCIAS

[1] J. Wu, K. Leung, G. Lau, *The Lancet*, 31 enero (2020)
 [2] J. Zheng, et al., *Applied Mathematics and Computation*, 162(2) (2005) 909-924
 [3] W. Guan et al., *The New England Journal of Medicine*, 28 febrero (2020)