Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Digital twins to fight against COVID-19 pandemic

Dongliang Chen, Nojoom A. AlNajem, Mohammad Shorfuzzaman

1. Introduction

Under the trend of rapid economic development, the exchanges among different regions have also deepened. The introduction of the “Belt and Road”, the increase of rural migrant workers, and the speed of travel tools such as high-speed rails and planes can promote the increase in exchanges among different cultures, but rapid movement of people causes that the number of patients infected with epidemic diseases (such as COVID-2019 (Corona Virus Disease 2019)) grows rapidly [1,2]. As a result, while intelligent technologies such as big data, Internet of Things (IoT), and deep learning continue to improve their performance, intelligent and effective prevention and control (P & C) of epidemics such as COVID-19, which is rapidly increasing infection speed, has become the focus of many research scholars in related fields.

Since the outbreak of the COVID-2019 in China on the eve of the Spring Festival in 2020, relevant government departments have adopted a series of measures such as extending the 2020 Spring Festival holiday, postponing the date of resumption of work, and delaying the opening of various schools. However, compared with the almost halted movement of people during the Spring Festival, the movement of people after the holiday still brings the risk of the spread of the epidemic [3]. Then, according to its transmission characteristics, prevention and control (P & C) measures such as home isolation for 14 days, reduction of the frequency of going out, and daily online reporting of health information, were adopted. During this period, all sectors of society, such as transportation, communications, key areas where personnel travel, public media, and Internet companies have all actively worked hard for the P & C of the epidemic. For example, Internet companies provide users with relevant information query and publishing services based on the data collected by themselves; railways and civil aviation collect the destination and contact information of passengers to provide follow-up tracking in the event of close contact; major operators provide users with path certification through mobile phone base station information; and major portal websites release the epidemic data and develop the health code [4]. These information collection and release services provided by the government or enterprises have provided a great help to the effective management and control of the epidemic, but they also have certain shortcomings.

This study is aimed to explore the anti-epidemic effect of artificial intelligence (AI) algorithms such as digital twins on the COVID-2019 (novel coronavirus disease 2019), so that the information security and prediction accuracy of epidemic prevention and control (P & C) in smart cities can be further improved. It addresses the problems in the current public affairs governance strategy for the outbreak of the COVID-2019 epidemic, and uses digital twins technology to map the epidemic P & C situation in the real space to the virtual space. Then, the blockchain technology and deep learning algorithms are introduced to construct a digital twins model of the COVID-2019 epidemic (the COVID-DT model) based on blockchain combined with BiLSTM (Bi-directional Long Short-Term Memory). In addition, performance of the constructed COVID-DT model is analyzed through simulation. Analysis of network data security transmission performance reveals that the constructed COVID-DT model shows a lower average delay, its data message delivery rate (DMDR) is basically stable at 80%, and the data message disclosure rate (DMDCR) is basically stable at about 10%. The analysis on network communication cost suggests that the cost of this study does not exceed 700 bytes, and the prediction error does not exceed 10%. Therefore, the COVID-DT model constructed shows high network security performance while ensuring low latency performance, enabling more efficient and accurate interaction of information, which can provide experimental basis for information security and development trends of epidemic P & C in smart cities.
For example, individuals or companies repeatedly submit information to different government departments, lacking of verification of the accuracy of the information provided, so that the companies and individuals can’t obtain reliable & C information of the epidemic in real time; and the effective privacy protection after the collection of citizens’ personal information is not complete. Therefore, the guarantee of information privacy during the P&C period of the epidemic and the effective prediction of the epidemic are of great significance to social development. With the ever-expanding trend of artificial intelligence (AI) applications such as blockchain, edge computing, and the IoT, its applications in the medical field are becoming more and more extensive. Among them, the distributed mechanism of blockchain technology achieves the selective assurance of privacy data security among participants and is anticipated to make manufacturing service transactions among makers more trustworthy [5]. Moreover, blockchain can ultimately improve the response to crisis events and realizes the secure sharing and automatic storage of private data [6]. The traceability of blockchain can enable the trajectory of the infected person to be tracked in time through the module based on geographic location [7,8]. However, due to the complexity of the resident trajectory in the real situation, the existing resident travel trajectory has to be mapped to the virtual space for analysis, that is, digital twins technology. The innovation and development of digital twins technology in the intelligent manufacturing industry has brought new guiding concepts to solve the problem of intelligent control of the intelligent epidemic P & C system driven by digital twins [9]. At the same time, deep learning is an algorithm for autonomously extracting data features, and applying it to the collection and prevention of medical data will make the development of the medical field more rapid. A reasonable neural network can approximate the nonlinear mapping with arbitrary accuracy, so as to perform systematic analysis and identification, pattern recognition, data compression, and intelligent detection of infectious diseases and other disease data [10, 11]. With the rapid progress in the medical field today, using the data feature extraction performance of deep learning algorithms will further accelerate the pace of disease-assisted diagnosis and increase the effect of accurate predictive analysis, so that the application of deep learning to the epidemic P and C has extremely important research value.

In summary, this study is developed to improve the intelligent efficiency and privacy security of epidemic P & C and ensure the general public can have a clearer understanding of their own health and epidemic prevention. The innovation of this study is to map the epidemic prevention and control situation in the real space to the virtual space through digital twins technology to reduce the interference factors; blockchain technology is introduced in response to security issues such as the disclosure of resident identity information during information sharing in the epidemic P & C process; and the BiLSTM (Bi-directional Long Short-Term Memory) in the deep learning algorithm is applied to accurately predict the P & C of the epidemic. Finally, simulation is performed to analyze its performance, which provides experimental basis for information security in later epidemic P & C process.

2. Related works

2.1. Application status of AI in network information security

With the fast growth of AI technology, the way of collecting epidemic information in the medical field has become more and more intelligent, and many scientific research scholars have conducted research on it. Alturjman et al. (2018) found that in medical applications, the 5G-inspired IIoT (Industrial Internet of Things) paradigm enables users to interact with secure WMSN (wireless medical sensor networks) with various types of sensors. However, the user’s self-verification before each interaction is a lengthy and time-consuming process, which can interfere with residents’ activities and reduce the overall performance of the medical system. In response to this problem, a CSIP (context-sensitive seamless identity provisioning) framework is proposed. Finally, it is found that the framework uses a secure mutual authentication method of hash and global assertion value to prove that the mechanism can achieve the main security goal of WMSN in a short time [12]. Kumar et al. (2018) used WSN (wireless sensor network) as a necessary platform for data perception and communication to describe the environment completely or make robust decisions. The spatiotemporal data mining is performed on the sensing samples collected by sensor nodes, and a redundancy elimination strategy is proposed to improve efficiency. Before forwarding to the base station or cluster head in the wireless sensor network, the collected data is mined to select the appropriate information. Finally, a large number of simulation experiments are performed, and the relevant results show that in the same simulation scenario, this program has better performance compared with other programs [13]. Khatkat et al. (2019) studied in detail the security issues of the perception layer of the IoT, and described its key components (ie, architecture, standards, and protocols) in the security environment of the perception layer. After the hierarchical security of the general IoT is described, two key enabling technologies of the perception layer, namely RFID (radio frequency identification devices) and sensor network, are discussed; and the open research issues and challenges related to the perception layer are identified and analyzed [14]. Muzammal et al. (2020) introduced several trust models based on the security requirements of IoT systems, and studied the security issues and requirements of IoT and RPL (routing protocol for low power and lossy networks), such as black holes, deception, and hierarchy. In addition, various mitigation methods and meanings of the trust model for secure routing in the IoT are also analyzed. Finally, they have a deeper understanding of trust measurement in the IoT environment, including open issues and research challenges, and the meaning of trust as a security paradigm for IoT networks and routing protocols [15]. Leng et al. (2021) proposed a novel iterative bi-level hybrid intelligence model named ManuChain, which can get rid of unbalance between holistic planning and local execution in individualized manufacturing systems [16]. Singh et al. (2021) analyzed the big data security and other related technologies using data collected by medical sensors in the health and health field and their new trends in solving real-world application challenges and investigated various well-known cryptography, biometrics, watermarking, and blockchain-based medical application security technologies [17].

2.2. Application status and trends of digital twins

Internet has grown rapidly, so that the scale of data in the medical field continues to grow, and the factors affected by the collected data are also increasing. For medical data processing and optimization, many scholars in related fields have conducted research on the impact of data collection in the medical field. Saddik (2018) applied digital twins to the construction of smart cities and found that digital twins help to monitor, understand, and optimize the functions of all physical entities, and provide continuous feedback to humans in smart cities to improve the quality of life and well-being [18]. Francisco et al. (2021) used smart meters to adjust building energy in smart cities digital twins, making the intelligent energy management within the geographic scope of large and medium-sized building becomes a key step in the construction of smart cities [19]. Laamarti et al. (2020) proposed an ISO/IEEE 11073 standardized digital dual-frame architecture, which includes the process of collecting data from personal health devices, analyzing these data, and delivering feedback to users in a cyclical manner. Finally, it is found that this framework can be used as a basis for the development of digital twins in smart medical care [20].

Digital twins can achieve semi-physical simulations to reduce the vast time and cost of physical commissioning and reconfiguration [21]. For instance, in the smart manufacturing system (SMS). In addition, the digital twin can directly conduct validation and test, which can quickly locate the malfunction and inefficiency reason, rule out the mistakes, and test the practicability of physical equipment in execution [22,23]. In response to the rapid development of intelligent transportation tools,
Zhang et al. (2021) proposed a vehicle edge caching mechanism based on social perception, which dynamically coordinates the caching capabilities of roadside units (RSUs) and the smart vehicles. At the same time, the digital twins technology is used to map the edge cache system into the virtual space to facilitate the construction of a social relationship model, and it is found through simulation that the constructed edge cache scheme has great advantages in optimizing cache utility [24].

To sum up, analysis of the above-mentioned research suggests that although scholars currently guarantee information security in the medical field, the information security and P & C predictions for the current global outbreak of the COVID-2019 epidemic are still not foolproof. Therefore, the blockchain technology is applied in this study to improve the security of the COVID-2019 epidemic information system, and deep learning and digital twins technology are adopted to predict and analyze the COVID-2019 epidemic. This is extremely important for the prevention and development of the global COVID-2019 epidemic.

3. Analysis of COVID-DT model based on blockchain and deep learning

3.1. Analysis on P & C and prediction demand of COVID-2019

At present, the epidemic P & C system in China mostly uses the platforms of the regional health departments of provinces and cities as the intermediate link, then uses AI technology to generate corresponding reports on the collected information, and finally uploads the information to the provincial and national epidemic network systems. Therefore, the epidemic intensity is judged to prevent the spread of the disease according to the early warning system of the epidemic in the system. However, there are still some shortcomings in the system, as shown in Fig. 1.

As shown in Fig. 1, the main problems existing in the current infectious disease P & C system are as follows. Firstly, the P & C of infectious diseases is not in place, so that the spread of infectious diseases is not effectively prevented. Secondly, there are hidden dangers in information security. Enterprises lack effective privacy protection after collecting citizens' personal information, which may disclose the private information of citizens. Thirdly, the distribution of medical supplies is unreasonable. During the outbreak of the COVID-19, the unreasonable distribution of materials among regions causes the infected persons to not be able to receive timely treatment, and the resources in the regions with abundant medical resources caused waste. Fourthly, the reporting procedure is onerous. There are many procedures and lack of multi-party verification at the same level, and medical institutions can't conduct in-depth communication on COVID-19. Fifthly, the real-time and traceability of data is insufficient. After the outbreak of COVID-2019, it is impossible to effectively track and monitor the source of COVID-19 virus in time, and the real-time delay is obvious [25,26]. This makes it impossible for relevant medical institutions to allocate protection resources reasonably during the outbreak of the COVID-19 epidemic, and ultimately leads to delays in real-time sharing of data resources.

Therefore, in the face of the current global outbreak of the COVID-2019, the protection of information security of people and the effective prevention and monitoring of epidemic are of extremely realistic significance for global economic development and social progress.

3.2. Analysis on safety P & C of blockchain technology on the COVID-DT model

When the virus breaks out and it is confirmed that the disease can be transmitted from person to person, the management department has to identify and isolate the patient, and track the people who have been in contact with the patient. In addition, it is necessary to ensure their information security while the information on the people in close contact is obtained. Therefore, it is very necessary to build an efficient and safe epidemic prevention system.

The blockchain itself also has security issues, which can be divided into three levels, namely process level, data level, and infrastructure level. There have been many studies on how different technologies can be
incorporated into blockchain to enhance the security, transparency, and traceability of the system [27,28]. However, the blockchain technology shows the characteristics of decentralization, openness, independence, security, and anonymity. Combined with the recent epidemic COVID-19 epidemic P & C hotspot, it is feasible in the sharing and updating of medical information and the construction of epidemic prevention and control system [29]. In response to the security issues such as the leakage of resident identity information in information sharing during the epidemic period, this study proposes to use the blockchain technology to protect the resident identity information in the network while guaranteeing the prevention of epidemic based on the traceability to travel information. The COVID-2019 epidemic P & C smart contract constructed using blockchain technology is shown in Fig. 2.

As shown in Fig. 2, in the smart contract architecture, the data contract participants are input from the outside, that is, the data and time of the COVID-2019 epidemic period, and the conditions for event response in the epidemic P & C are preset. The state machine and contract transaction set are used to execute the current state of the network and select the execution of the next contract transaction, and finally achieve the docking with the external epidemic P & C information. At the same time, it can judge whether the trigger condition of the contract is satisfied according to the preset response rules, and the relevant commands are automatically executed by the computer network. Among them, the Hash algorithm is undertaken as the core content of blockchain technology, which can map message m of any unit length into output value H(m) of fixed unit length. H(m) refers to the message digest of epidemic data [30], which can be expressed as follows:

\[ h = H(m) \]  

If the same m is inputted, the output value h is always the same. The Hash algorithm is public and shows three important characteristics: one-way, easy to calculate, and collision resistance. With the further popularization of information technology, the digital signature scheme for the P & C system of the epidemic using blockchain technology is shown in Fig. 3.

When digitally signing the epidemic P & C system is performed, the collected personnel epidemic prevention data information format has to be packaged, which is recorded as m. In the process of signing the prevention message, the end user who has the right to forward sends the epidemic prevention data message m to the network edge node. The first terminal user sends the epidemic prevention message m to other user terminals on this link for signature, and other user terminals on this link respond to the epidemic P & C message m sent by the first collected user terminal. The signing process is as follows. Step 1: the other user terminal is allowed to select a random number \( k \in [1, n - 1] \) and calculate the point \((x_1, y_1)\) according to the following equation:

\[ (x_1, y_1) = kG(x, y) \]  

**Step 2.** \( r \) is calculated according to the equation (3). If \( r = 0 \), it has to return to step 1 to reselect the random number \( k \):

\[ r = x_1 \mod n \]  

**Step 3.** If \( r \neq 0 \), it can calculate the hash value of the epidemic P & C message \( m \) according to equation (4):

\[ h(m) = SHA256(m) \]  

**Step 4.** \( S_r \) is calculated based on the random number \( k \), the hash value \( h(m) \) of the epidemic P & C message, the private keys of other user terminals \( SK_v \), and the equation (5). If \( S_r = 0 \), it has to return to the step 1 to reselect the random number \( k \).

\[ S_r = [(h(m) + SK_v \cdot r) \cdot k^{-1}] \mod n \]  

**Step 5.** If \( S_r \neq 0 \), the signature of the hash value \( h(m) \) of the epidemic P & C message by other user terminals is \( \{r, S_r\} \).

It should further verify and sign the end user information in the P & C system of the epidemic. The information received by the network edge node includes the road condition message \( m \) sent by the first end user and the signature \( \{r, S_r\} \) of the hash value \( h(m) \) of the epidemic P & C message by other end users. The specific process of network edge nodes verifying and signing epidemic P & C messages is as follows:

In the first step, if \( \{r, S_r\} \) is an integer within \([1,n-1]\), \( h(m) \) of the epidemic P & C message \( m \) can be calculated from the first end user to receive the epidemic P & C message according to equation (6). Secondly, \( X \) can be calculated based on \( h(m) \) of the epidemic P & C message \( m \) sent by the first end user calculated by the network edge node, the signature \( \{r, S_r\} \) of the \( h(m) \) of the epidemic P & C warning message by other end users, the public key \( PK_v \) of other end users, and equation (6) below.

\[ X = [h(m) + S_r \cdot r \cdot G(x, y) + r \cdot S_r^{-1} \cdot PK_v] \mod n \]  

In the third step, if \( X = 0 \) or \( \infty \), the signature is rejected, otherwise, the projection of \( X \) on the x-axis in the rectangular coordinate system is calculated and recorded as \( x' \). In the fourth step, \( f \) is calculated according to equation (7), if \( f = r \), then the network edge node passes verification.

\[ f = x' \mod n \]  

When the network edge node collects the epidemic P & C messages sent by other end users on this link, this study uses pseudonym exchange to ensure that the identity of legitimate end users is not obtained by attackers so as to ensure its privacy and improved security performance. After the registration of \( v_i \) is completed, the trust center TA can use the real identity \( RID_i \) of \( v_i \) to generate a pseudonym for it. That is, TA chooses a random number \( w_i \in Z_q \) and uses the below equation to calculate the pseudonym:

\[ RID_{i1} = w_i \cdot g \]  

\[ RID_{i2} = RID_{i1} \oplus H(w_i \cdot PK_i) \]  

From this, TA can get the initial pseudonym \( RID = \{RID_{i1}, RID_{i2}\} \), which is stored in the tracking list, and sent to \( v_i \) via the network edge node. Then, the \( RID \) of \( v_i \) can be obtained through its own private key, which is not obtained by other end users. Furthermore, exchange entropy is used to measure the strength of privacy information protection of end
users, and exchange entropy calculation is used to control pseudonym exchange conditions [31]. Maybe there is an anonymous set $C = \{v_1, v_2, \ldots, v_n\}$ of terminal users in a pseudonym exchange area, which means that a total of $m$ terminal users may exchange pseudonyms with each other, and the exchange entropy of $v_i$ is expressed by equation (10):

$$E_i = -\log \pi_i$$

(10)

In the equation above, $\pi_i$ refers to the probability that the terminal user is tracked after the pseudonym is exchanged once, and the greater the entropy, the higher the information privacy of the terminal user. The exchange entropy of the end user set $C$ is given as follows:

$$E_C = -\sum_{i=1}^{m} \pi_i \log \pi_i$$

(11)

$P_s(v_i|C_{ik})$ is adopted to represent the probability of $v_i$ to select an internal opponent for pseudonym exchange, and the expression is as follows:

$$P_s(v_i|C_{ik}) = \frac{A}{N - 1}$$

(12)

In the above equation, $N$ refers to the total end users in the current area, and $A$ refers to the number of internal opponents. Then, the exchange entropy increased by the end user after the pseudonym exchange is written as follows:

$$\Delta E = \sum_{i=1}^{k} \left( \frac{A}{N - 1} \right) \frac{(N - 1) - i}{(N - m - 1)} \log_2(m - i)$$

(13)

Once $v_i$ is exchanged with the attacker during the pseudonym exchange, the exchange entropy after this exchange is 0; conversely, if $v_i$ does not encounter the attacker in the $m$-th pseudonym exchange, then the exchange entropy after the exchange is shown in equation (14):

$$E'_i(v_i \notin C_{ik}) = E_i(m - 1) + \Delta E$$

(14)

The epidemic P & C system takes equation (14) as the basis for measuring the benefits and risks of pseudonym exchange, and undertakes equation (15) to express the conditions for end users to participate in the pseudonym exchange:

$$\Delta E > \frac{P_s(v_i|C_{ik})E_i^{n-1}(v_i \notin C_{nk})}{1 - P_s(v_i|C_{ik})}$$

(15)

End users who meet the above constraints can be referred to as intimate end users.

Among the security factors that affect the P & C of the epidemic, in addition to the privacy and security of end users, the security of epidemic information is another key factor. This study is developed based on Shamir’s $(k, n)$ secret sharing method to protect the message security to improve the security and reliability of the system. Specifically, the secret message $M_i$ to be shared is divided into $n$ fragment messages by $v_i$ according to the situation of the end user. In the process of message fragmentation, when the source end user $v_i$ determines the source message $M_i$ and the number of fragment messages, the $h(m)$ is firstly calculated, where $M_i$ and $h(M_i)$ are two independent secret data [32]. Then, according to the equation (16) below, $v_i$ can realize the generation and distribution of the message:

$$f(x_i) = M_i + a_1x_i + a_2x_i^2 + \cdots + a_{n-1}x_i^{k-1}$$

(16)

In the equation (16) above, $a_1, \ldots, a_{n-1}$ is the random number generated by $v_i$, and $\pi_i$ refers to the secret sharing threshold, that is, the minimum number of fragment messages required to restore the message. The security of the $(k, n)$ threshold is to prevent less than $k$ end users from cooperating to recover the secret, and to prevent attacks on the P & C of the epidemic from hindering the recovery of the secret. The value of $k = t = (n - 1)/2$ is the best at $n = 2t + 1$. Then, to ensure the security of $M_i$, the Lagrangian difference algorithm is used to complete the recovery of the fragments. It is assumed that any message fragment $(j, f(j))$ of $M_i$ has a $j = 0, 1, \ldots, k - 1$, the k polynomials $f_i(x) = \prod_{i \neq k}^{x} \prod_{x}^{x_i}$ can be defined as below equation according to the Lagrangian difference algorithm:

$$f_i(x) = \prod_{i \neq k}^{x} \prod_{x}^{x_i}$$

(17)

In the above equation, $f_i(x)$ is a n-1 polynomial, $l = 1, 2, \ldots, n$, $I_i = \{1, 2, \ldots, l, \ldots, k\}$, and $\forall i \in I_i$ is satisfied, then below equation can be satisfied.

$$f(x) = \sum_{i=1}^{k} f_i(x)$$

(18)

Thus, the source message is recovered as $M'_i = f(0)$, and then $h(M'_i)$ is calculated and is compared with $h(M_i)$ to obtain the recovery conditions:

$$h(M_i) = h(M'_i)$$

(19)

If equation (19) is not satisfied, it indicates that the restoration has failed; otherwise, it indicates that the restoration is successful. Finally, the network edge node sends the source message to the end users of the epidemic P & C system in the form of broadcast or graph.

To further predict the P & C situation of the epidemic P & C system, it introduces a deep learning algorithm in this study. As a variant of cyclic neural network, Bi-LSTM network has been applied in many fields, and it has obtained great achievements in natural language processing, text processing, and time series data prediction [33]. The basic framework is shown in Fig. 4.

As shown in Fig. 4, two LSTMs are used in the hidden layer to model the sequence from front to back (forward) and back to front (reverse), and then their outputs are connected [34]. Among them, the RNN calculation equation from front to back is given in equation (20), and the RNN calculation from back to front is shown in equation (21):

$$h'_i = H(W^i - [h'_{i-1}, x_i] + b^i)$$

(20)
$h_t = H(W^r \cdot [h_{t-1}, x_t] + b^r)h_t = [h_t^r : h_t^c]$  

(21)

$h_t = [h_t^r : h_t^c]$ at time $t$ means to join $h_t^r$ and $h_t^c$ end to end. BiLSTM replaces the cyclic unit in BiRNN with an LSTM structure, so that the advantages of both BiRNN and LSTM can be combined to identify the P & C of COVID-19.

The fluctuations in the incidence of infectious diseases such as COVID-2019 are affected by changes in natural, social, and economic factors, and the change process of these three factors is slow and regular. Therefore, the incidence of infectious diseases can be predicted.

3.3. Analysis on COVID-DT model based on blockchain combined with BiLSTM

In response to the shortcomings in the public affairs governance strategy of the current COVID-2019 epidemic, the epidemic P & C situation in the real space is mapped to the virtual space in this study; then the blockchain technology is introduced and the consensus mechanism is used to ensure that the accounting copies maintained by all nodes on the blockchain are synchronized with each other. In addition, the legal transactions are added to the blockchain after negotiation. The consensus...
mechanism is the core part of the safe operation of the blockchain, which can ensure the safety of medical information sharing while improving the sharing efficiency, playing an irreplaceable role. To further predict the P & C situation of the epidemic P & C system, the Bi-LSTM network in the deep learning algorithm is introduced to predict the P & C of the COVID-19 epidemic, and a COVID-DT model is constructed based on blockchain combined with BiLSTM, as shown in Fig. 5.

In this epidemic P & C model, \( v_i \) has to consider the reputation of the other party's past behavior from direct trust and indirect trust to further reduce the risk of sending messages to malicious network edge nodes. The direct trust degree refers to the direct trust level estimated by the source data of the end user \( v_i \) and the direct node end user \( v_j \) based on the history of direct interaction behavior (the situation of sending and receiving data packets) within a period of time \([35]\).

The direct trust degree refers to the direct trust level estimated by the source data of the end user \( v_i \) and the direct node end user \( v_j \) based on the history of direct interaction behavior (the situation of sending and receiving data packets) within a period of time \([35]\). \( g(p) \) is assumed to be the prior probability and the loss is the square loss, the characteristic that the posterior probability distribution of the binomial event obeys the Beta distribution can be used to derive the Bayesian estimation of the trust relationship \( d'(x) \) can be written as equation below:

\[
d'(x) = E[p|X=x] = \int_0^1 pg(p|x)dp
\]  

In the equation (22) above, \( g(p|x) \) refers to the posterior probability density of \( p \), which can be expressed by the following equation:

\[
g(p|x) = \frac{\sum_{i=1}^n x_i}{\Psi(n+1)} (1-p)^{n-\sum_{i=1}^n x_i} (p)^{\sum_{i=1}^n x_i}
\]  

\( \Psi(a, b) \) in the equation (23) can be expressed as equation (24):

\[
\Psi(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}
\]  

Then, the equation (23) can be transferred into below equation:

\[
g(p|x) = \frac{\sum_{i=1}^n x_i}{\left(\sum_{i=1}^n x_i\right)!} (1-p)^{n-\sum_{i=1}^n x_i} (p)^{\sum_{i=1}^n x_i}
\]  

Finally, the Bayesian estimation function \( d'(x) \) of the trust relationship can be expressed as follows:

\[
d'(x) = \frac{\sum_{i=1}^n x_i}{\left(\sum_{i=1}^n x_i\right)!} (1-p)^{n-\sum_{i=1}^n x_i} (p)^{\sum_{i=1}^n x_i} \int_0^1 dp
\]  

According to the above Bayesian estimation, the direct trust value is expressed in the below equation:

\[
D_i(w, t) = \frac{w + 1}{t + 2}
\]  

In the equation (27) above, \( t \) refers to the number of verifications, and \( w \) represents the number of subsequent trust successes. If there are \( x \) types of direct experiences to be verified between the source terminal users \( v_i \) and \( v_j \), then when the direct trust value of \( v_j \) is calculated, different types of experiences have to be assigned with corresponding weights \( w_n \), and then the direct trust value of the source terminal user \( v_i \) and \( v_j \) can be calculated with equation (28):

\[
DT_i = \sum_{n=1}^{x} W_n D_i(w, t)
\]  

In the indirect trust evaluation, when the trust evaluation is performed for \( v_i \) and \( v_j \), the recommendation trust degree of the recommender \( v_i \) is \( R_{ij} \), which can be calculated as follows:

\[
R_{ij} = \frac{\sum_{w=\text{CN}(i, r)} (\sum_{w=CN(i, r)} (DT_{is} - DT_i) \times (DT_{is} - DT_i))}{\sqrt{\sum_{w=\text{CN}(i, r)} (DT_{is} - DT_i)^2} \times \sqrt{\sum_{w=\text{CN}(i, r)} (DT_{is} - DT_i)^2}}
\]  

In the equation (29), \( \text{CN}(i, r) \) refers to the common neighbor node between the node \( v_i \) and the recommender \( v_j \), and the direct trust degrees of \( v_i \) and \( v_j \) to \( v_j \) are \( DT_{is} \) and \( DT_{js} \), respectively; \( DT_i \) and \( DT_j \) are the average direct trust degrees of \( v_i \) and \( v_j \), respectively. The below expression equation can be obtained by observing the interaction between the end users \( v_i \) and \( v_j \) and \( \text{CN}(i, r) \):

\[
\begin{align*}
DT_i &= \frac{1}{n} \sum_{w=1}^{n} DT_i \\
DT_j &= \frac{1}{n} \sum_{w=1}^{n} DT_j \\
\end{align*}
\]  

Eventually, the P & C message of the epidemic is transmitted along the direction of increasing trust gradient, and the message is forwarded only when \( T_i > T_j \). The possibility of messages reaching the server gradually increases, and the success rate of message forwarding is also improved. The attacker’s trust is very low, and the reason is that the attacker refuses to communicate and cooperate. Therefore, the attacker
can be excluded based on the trust evaluation, thereby ensuring the security and reliability of the network.

To further predict the development of the epidemic in the epidemic P & C system, it has to pre-process the data information firstly. Due to the input and output ranges of the nonlinear activation function used in the neural network prediction model, it is necessary to normalize the variable data of the four dimensions to unify the dimension to facilitate the calculation of the gradient, accelerate the convergence, and avoid neuron saturation. The normalization standard used in this study is the minimum and maximum standardization method, and the corresponding equation is as follows:

$$ P_0 = \frac{P}{P_{\text{max}}} \frac{P_{\text{min}}}{P_{\text{max}}} $$

(31)

$P_0$ refers to the original data value at the current moment; $P_{\text{min}}$ and $P_{\text{max}}$ refer to the minimum and maximum values of all data, respectively; and $P'$ is the normalized result of all end-user source data sought.

After the normalization is completed in the data input stage, the actual output predicted data has to be restored to its original size by using a normalized reduction function. The reduction function corresponding to the minimum and maximum standardization method is given as follows:

$$ P_{\text{pre}} = P_{\text{pre}}' \frac{P_{\text{pre}}}{P_{\text{max}}} + P_{\text{min}} $$

(32)

$P_{\text{pre}}'$ and $P_{\text{pre}}$ in the above equation refer to the output value predicted by the unreduced Bi-LSTM and the P & C value COVID-19 epidemic predicted by the reduced Bi-LSTM, respectively. The specific prediction process of the COVID-DT model based on the neural network algorithm is shown in Fig. 6.

### 3.4. Simulation analysis

A simulation experiment can conduct to verify the effectiveness of the proposed approach to achieving improved system performance while avoiding the vast cost of physical commissioning of the manufacturing system [36]. To verify the performance of the model constructed, a COVID-2019 epidemic digital twins system is built on the Matlab network simulation platform based on hardware and software design. The blockchain technology uses the Hyperledger Fabric alliance blockchain platform, and uses tensorflow to build the neural network in the model. The experimental data comes from the number of new crown epidemic cases in Shandong province from February 2020 to May 2020. After the obtained data is pre-processed, it is divided into training set and test set according to the ratio of 8:2. In the simulation, the CPU (Central Processing Unit) model of the computer is CORE-i7-4720HQ-2.6 GHz, and the neural network is built using open source Tensorflow framework of Google, which is a machine learning and deep learning programming framework based on vector flow graphs. Matrix calculations are completed using Numpy and Pandas open source toolkits. Numpy library is an open source matrix processing library. Chaincode (smart contract) is implemented in Go language. Pandas library provides good help for data cleaning and data preprocessing in data analysis.

The blockchain-based BiLSTM algorithm proposed in this study is compared with CFT (Crash Fault Tolerance) [37], BFT (Byzantine Fault Tolerance) [38], PBFT (Practical Byzantine Fault Tolerance) [39], HPBFT (Hybrid Parallel Byzantine Fault Tolerance) [40], and DAPP (Delay-aware and privacy-preserving) [41] from the perspectives of the number of end users, the survival time of different data messages, throughput, and cost, so as to analyze the safety performance of COVID-DT model constructed. In addition, the accuracy in predicting the COVID-2019 epidemic P & C is analyzed from the perspective of root mean square error (RMSE), average absolute error (MAE), and average absolute
Fig. 8. Comparison on network data security transmission of each mechanism algorithm under different data message survival times (a: average DMDR; b: average DMDCR; c: average delay).

Fig. 9. Trend change curves of communication cost under different mechanism algorithms (a: communication cost; b: verification cost; c: storage cost).
percentage error (MAPE). In order for the neural network framework to achieve the desired prediction results, the following hyperparameters need to be set. The number of training cycles (Epochs) is 30, the learning rate is 0.002, the batch size is 128, the activation function is ReLU, the dropout rate is 0.5, and the optimizer is Adam.

4. Results and discussion

4.1. Performance of network data security transmission of the COVID-2019 epidemic based on different mechanisms and algorithms

The COVID-DT model constructed is compared with CFT, BFT, PBFT, HPBFT, and DAPP from the perspectives of the number of end users, the survival time of different data messages, throughput, and cost, so as to analyze its safety performance. The specific results are given in Fig. 7~10.

The network data security transmission performance of each model algorithm under different terminal user numbers is compared, and the results are shown in Fig. 7. With the increase in the number of end users, the average DMDR shows an upward trend, and that of the COVID-DT model is not less than 80% (Fig. 7a). The average DMDCR has basically not changed greatly, and that of the COVID-DT model does not exceed 10% (Fig. 7b). As the number of end users increases, the average delay decreases, and the average delay of autonomous vehicle data transmission in COVID-DT model is basically stable at about 350 ms (Fig. 7c). Therefore, from the perspective of the number of different end users, the model algorithm constructed clearly shows that the average DMDR is higher, the average DMDCR is the lowest, and the delay is lower, thus showing good network data security transmission performance.

The network data security transmission performance of each mechanism algorithm under different data message survival times is compared, and the results are shown in Fig. 8. As the survival time of data messages increases, the model used in this study shows a better average DMDR and a lower DMDCR. This may be due to the fact that the model
uses message fragmentation and the forwarding strategy of trusted end users, while other algorithms use message fragmentation to protect the message security, lacking the trust evaluation of the forwarding node. As the survival time of epidemic P & C data messages increases, the average delay gradually increases, and the average delay of the model proposed is basically stable at about 150 ms.

Fig. 9 illustrates the analysis results of the communication cost of each mechanism algorithm. Compared with other solutions, the communication cost of sending messages will be lower in this scheme, and the corresponding communication cost gap will become more and more obvious as the messages sent increase (Fig. 9a). As the number of verification signatures increases, the verification cost in this study is shorter, which makes the delay time shorter, while the delay of other schemes increases linearly (Fig. 9b). As the number of end users increases, the storage cost still appears to be the lowest, and does not exceed 700 bytes (Fig. 9e). Therefore, this solution can make information interaction more efficient and have a lower delay in contrast to other algorithms.

As shown in Fig. 10, the throughput increases with the number of end users under different protocol algorithms. Fig. 10a shows the receiving throughput, which indicates that algorithm adopted in this study is the highest, followed by the HPBFT, and that of DAPP is the lowest. Receiving throughput refers to the number of packets received from its neighbors. Higher receiving throughput means that end users can obtain more information about their neighbors, which can meet the requirements of end users’ micro perception. Sending throughput refers to how many data packets can be sent per frame per user. Due to the requirements of safety-related applications, the transmission throughput should be approximately one packet per frame to ensure that each end user has the opportunity to obtain its location, local epidemic P & C status, and so on. The simulation shows that, except for the DAPP protocol algorithm, others algorithms have almost the same sending throughput (1 data packet per user per frame), which is close to 1, and the algorithm in this research is the closest to 1. However, as the number of vehicles increases, the transmission throughput of DAPP decreases. It directly leads to the inability of some end users to obtain time slots and can’t meet the demanding requirements of security applications (Fig. 10b). Therefore, the COVID-DT model based on blockchain combined with BiLSTM can adaptively change the communication range, resulting in less interference and increasing network throughput.

4.2. Analysis on epidemic P & C prediction performance of each mechanism algorithm

The system model constructed is compared with other mechanism algorithms from RSME, MAE, and MAPE to analyze the prediction performance of COVID-DT model, and the specific comparison results are shown in Fig. 11.

Fig. 11 illustrates the epidemic prediction error of each mechanism algorithm. The blockchain combined with the BiLSTM algorithm proposed is compared with the mechanism algorithm used by other scholars. Under the COVID-2019 epidemic P & C prediction, the RSME, MAE, and MAPE are 4.95%, 6.62%, and 8.74%, respectively, while those of other model algorithms are higher obviously. Therefore, the COVID-DT model based on blockchain combined with BiLSTM can greatly reduce the prediction error of the COVID-2019 epidemic data in various scenarios in the city, achieving a more accurate prediction effect.

5. Conclusion

With the rapid growth of science and technology, technologies such as digital twins and blockchain have shown great application potential. Aiming at the public health complex situation of the COVID-2019 epidemic in real space, this study introduces digital twins technology, maps the epidemic P & C in the real network to the virtual space, combines the blockchain technology and deep learning algorithms, and builds a COVID-DT model based on blockchain combined with BiLSTM. Finally, the simulation experiment reveals that the model algorithm proposed shows a better average DMDR (about 80%), a lower DMDCR (below 20%), and the lowest delay (not more than 200 ms); and the communication cost was not more than 700 byte. Therefore, it can realize more efficient and accurate data information interaction in the epidemic data transmission network, and it can provide experimental basis for the information security in the epidemic P & C. However, there are some shortcomings. Although it has completed the research and implementation of the COVID-DT model based on digital twins simulation modeling, more work is still needed to further realize the application of digital twins. Firstly, the current control system is the one used in the Kebo Company, and the system can’t be fully opened. It will use a domestic system for further analysis in the next step. Secondly, the application of big data analysis technology to continue to improve the “service system” of digital twins technology is extremely important for the subsequent development of P & C and precise prediction of the COVID-2019 epidemic.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China (No. 61902203).

References

[1] J. Pang, Y. Huang, Z. Xie, J. Li, Z. Cai, Collaborative city digital twin for the COVID-19 pandemic: a federated learning solution, Tsinghua Sci. Technol. 26 (5) (2021) 759–771.
[2] R.G. Díaz, F. Laamarti, A. El Saddik, DTCoach: your digital twin coach on the edge during COVID-19 and beyond, IEEE Instrum. Meas. Mag. 24 (6) (2021) 22–28.
[3] O. Okorie, R. Subramaniam, F. Charlassy, J. Patavellas, D. Widdifield, K. Salomits, Manufacturing in the time of COVID-19: an assessment of barriers and enablers, IEEE Eng. Manag. Rev. 48 (3) (2020) 167–175.
[4] M. Mashaly, Connecting the twins: a review on digital twin technology & its networking requirements, Procedia Comput. Sci. 184 (2021) 299–305.
[5] J. Leng, P. Jiang, K. Xu, Q. Liu, J.L. Zhao, Y. Bian, R. Shi, Makerchain: a blockchain with chemical signature for self-organizing process in social manufacturing, J. Clean. Prod. 234 (2019) 767–778.
[6] J. Leng, S. Ye, M. Zhou, J.L. Zhao, Q. Liu, W. Guo, L. Fu, Blockchain-secured smart manufacturing in industry 4.0: a survey, IEEE Trans. Syst. Man Cybernet.: Systems 51 (1) (2020) 237–252.
[7] Q. Lin, H. Wang, X. Pei, J. Wang, Food safety traceability system based on blockchain and EPCIS, IEEE Access 7 (2019) 20698–20707.
[8] C. Roy, S. Misra, S. Pal, Blockchain-enabled safety-as-a-service for industrial IoT applications, IEEE Internet Things Mag. 3 (2) (2020) 19–23.
[9] D. Ivanov, Digital supply chain management and technology to enhance resilience by building and using end-to-end visibility during the COVID-19 pandemic, IEEE Trans. Eng. Manag. (2021) 1–11.
[10] S. Hu, Y. Gao, Z. Niu, Y. Jiang, L. Li, X. Xiao, G. Yang, Weakly supervised deep learning for covid-19 infection detection and classification from ct images, IEEE Access 8 (2020) 118869–118883.
[11] R. Kumar, A.A. Khan, J. Kumar, A. Zakria, N.A. Goliirza, S. Zhang, W. Wang, Blockchain-federated-learning and deep learning models for covid-19 detection using ct imaging, IEEE Sensor. J. 21 (14) (2021) 16301–16314.
[12] F. Al-Turjuman, S. Alturjuman, Context-sensitive access in industrial internet of things (IIoT) healthcare applications, IEEE Trans. Ind. Inf. 14 (6) (2018) 2736–2744.
[13] S. Kumar, V.K. Chaurasiya, A strategy for elimination of data redundancy in internet of things (IoT) based wireless sensor network (WSN), IEEE Syst. J. 13 (2) (2018) 1650–1657.
[14] H.A. Khattak, M.A. Shah, S. Khan, I. Ali, M. Imran, Perception layer security in internet of things, Future Generat. Comput. Syst. 100 (2019) 144–164.
[15] S.M. Muzammal, R.K. Murugesan, N.Z. Jhanjhi, A comprehensive review on secure routing in internet of things: mitigation methods and trust-based approaches, IEEE Internet Things J. 8 (6) (2020) 4186–4210.
[16] J. Leng, D. Yan, Q. Liu, K. Xu, J.L. Zhao, R. Shi, X. Chen, ManuChain: combining permissioned blockchain with a holistic optimization model as bi-level intelligence for smart manufacturing, IEEE Trans. Syst. Man Cybernet.: Systems 50 (1) (2019) 182–192.
[17] A.K. Singh, A. Anand, Z. Li, K. Ko, A. Mohan, A survey on healthcare data: a security perspective, ACM Trans. Multimed. Comput. Commun. Appl. 17 (2) (2021) 1–26.
A. Saddik, Digital twins: the convergence of multimedia technologies, IEEE Multimedia 25 (2) (2018) 87–92.

A. Francisco, N. Mohammadi, J.E. Taylor, Smart city digital twin-enabled energy management: toward real-time urban building energy benchmarking, J. Manag. Eng. 36 (2) (2020), 04019045.

F. Laamarti, H.F. Badawi, Y. Ding, F. Arafsha, A. El Saddik, An ISO/IEEE 11073 standardized digital twin framework for health and well-being in smart cities, IEEE Access 8 (2020) 105950–105961.

J. Leng, M. Zhou, Y. Xiao, H. Zhang, Q. Liu, W. Shen, L. Li, Digital twins-based remote semi-physical commissioning of flow-type smart manufacturing systems, J. Clean. Prod. 306 (2021), 127278.

J. Leng, D. Yan, D. Zhang, L. Wei, A. Yu, X. Chen, Digital twin-based design of the configuration, motion, control, and optimization model of a flow-type smart manufacturing system, J. Manuf. Syst. 58 (2021) 52–64.

J. Leng, D. Yan, Q. Liu, H. Zhang, G. Zhao, L. Wei, X. Chen, Digital twin-driven joint optimisation of packing and storage assignment in large-scale automated high-rise warehouse product-service system, Int. J. Comput. Integrated Manuf. 34 (7–8) (2021) 783–800.

K. Zhang, J. Cao, S. Maharjan, Y. Zhang, Digital twin empowered content caching in social-aware vehicular edge networks, IEEE Trans. Comput. Soc. Syst. (2021) 1–13.

S. Bonadies, S.A. Gadsden, An overview of autonomous crop row navigation strategies for unmanned ground vehicles, Eng. Agril. Environ. Food 12 (1) (2019) 24–31.

M.H.F.M. Fauadi, S. Akmal, M.M. Ali, N.I. Anuar, S. Ramlan, A.Z.M. Noor, N. Awang, Intelligent vision-based navigation system for mobile robot: a technological review, Period. Eng. Nat. Sci. 6 (2) (2018) 47–57.

J. Leng, M. Zhou, J.L. Zhao, Y. Huang, Y. Bian, Blockchain Security: A Survey of Techniques and Research Directions, IEEE Transactions on Services Computing, 2020.

J. Leng, G. Ruan, P. Jiang, K. Xu, Q. Liu, X. Zhou, C. Liu, Blockchain-empowered sustainable manufacturing and product lifecycle management in industry 4.0: a survey, Renew. Sustain. Energy Rev. 132 (2020), 110112.

A. Gu, Z. Yin, C. Cui, Y. Li, Integrated functional safety and security diagnosis mechanism of GFS based on blockchain, IEEE Access 8 (2020) 15241–15255.

J. Chen, W. Wang, Y. Zhou, S.H. Ahmed, W. Wei, Exploiting 5G and blockchain for medical applications of drones, IEEE Network 35 (1) (2021) 30–36.

X. Zhang, P. Sun, J. Xu, X. Wang, J. Yu, Z. Zhao, Y. Dong, Blockchain-based safety management system for the grain supply chain, IEEE Access 8 (2020) 36398–36410.

A. Lei, H. Cnuckshank, Y. Cao, P. Asuquo, C.P.A. Ogah, Z. Sun, Blockchain-based dynamic key management for heterogeneous intelligent transportation systems, IEEE Internet Things J. 4 (6) (2017) 1832–1843.

M. Jamshidi, A. Lalbakhsh, J. Talha, Z. Peroutka, F. Hadjiboei, P. Lalbakhsh, W. Mohyaddin, Artificial intelligence and COVID-19: deep learning approaches for diagnosis and treatment, IEEE Access 8 (2020) 109581–109595.

M.A. Rahman, M.S. Hossain, N.A. Alrajeh, N. Guizani, BSG and explainable deep learning assisted healthcare vertical at the edge: COVID-19 perspective, IEEE Network 34 (4) (2020) 98–105.

Q. Tao, X. Cui, X. Huang, A.M. Leigh, H. Gu, Food safety supervision system based on hierarchical multi-domain blockchain network, IEEE Access 7 (2019) 51817–51826.

J. Leng, Q. Liu, S. Ye, J. Jing, Y. Wang, C. Zhang, X. Chen, Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model, Robot. Comput. Integrated Manuf. 63 (2020), 101895.

H. Xu, L. Zhang, Y. Liu, B. Cao, RAFT based wireless blockchain networks in the presence of malicious jamming, IEEE Wireless Commun. Lett. 9 (6) (2020) 817–821.

Y. Li, L. Qiao, Z. Lv, An optimized byzantine fault tolerance algorithm for consortium blockchain, Peer-to-Peer Network. Appl. (2021) 1–14.

S. Gao, T. Yu, J. Zhu, W. Cai, T-PBFT: an EigenTrust-based practical Byzantine fault tolerance consensus algorithm, China Commun. 16 (12) (2019) 111–123.

X. Xu, D. Zhu, X. Yang, S. Wang, L. Qi, W. Dou, Concurrent practical byzantine fault tolerance for integration of blockchain and supply chain, ACM Trans. Internet Technol. 21 (1) (2021) 1–17.

G. Xu, H. Li, S. Liu, M. Wen, R. Lu, Efficient and privacy-preserving truth discovery in mobile crowd sensing systems, IEEE Trans. Veh. Technol. 68 (4) (2019) 3854–3865.