

CYBER SECURITY OF MULTI-CLOUD HEALTHCARE SYSTEMS: A HIERARCHICAL DEEP LEARNING APPROACH

***Rashmi. K¹, Marie Claude. N², Nandha Gopal. S. M³, Rejina Parvin. J⁴, Roopa. B. S⁵, Nikita Katariya⁶,
Ramana Reddy. B⁷**

¹Assistant Professor, Department of Computer Science and Engineering (AIML), Vidyavardhaka College of Engineering, Mysore, Karnataka, India, rashmim.k@vvce.ac.in

²Associate Professor, Chennai Institute of Technology, Chennai, Tamilnadu, India,
drmarieclauden@gmail.com

³Professor, Department of Computer Science and Engineering, ACS College of Engineering, Bangalore, Karnataka, India, nandhagopalsm@gmail.com

⁴Professor, Department of CSE – Cyber Security, Sri Krishna College of Engineering & Technology, Coimbatore, Tamilnadu, India., rejinaparvinj@skcet.ac.in

⁵Assistant Professor, Department of Electronics and Communication Engineering, Jawaharlal Nehru New College of Engineering, Shimoga, Karnataka, India, roopabs92@gmail.com

⁶Assistant Professor, School of Computer Science and Engineering, Shri Ramdeobaba College of Engineering and Management, Ramdeobaba University, Nagpur, Maharashtra, India,
katariyanikita08@gmail.com

⁷Assistant Professor, Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India, brreddycse@gmail.com

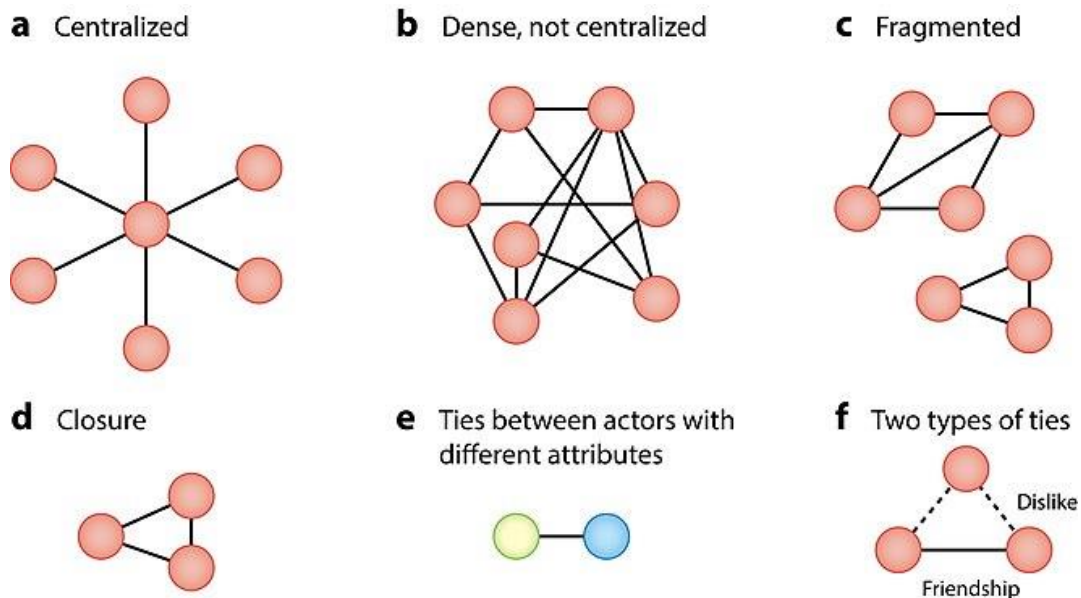
Abstract

This article describes the architecture of a typical mobile health application. Anyone using the system can set their own personalized privacy level. This article describes how social networking communications can be made more secure and private in a multi-cloud environment, especially for medical applications. The purpose of this study is to present a practical approach for extracting individual-level AH from GSV data. We have developed an entirely new hierarchical deep learning method that can identify AH based on global and local visual features. The objective of this study was to present a practical technique for separating individual-level BP from GSV data. To determine AH using global and local visualization features, we created a new hierarchical deep learning system. These systems become more vulnerable to hacking as technology advances. The research presented in this white paper aims to show the feasibility and utility of artificial intelligence-based hierarchical deep neural network models to protect data in motion. A new aspect of this research is the synergistic formation of multiple clouds using different complex models, which shortens the formation

time and increases the accuracy of detecting unidentified attacks. receive. As a result, the core cloud model converges much less time (6-8) and far fewer epochs than the boundary cloud model, which takes 35-40 epochs to train. Extensive evaluation has demonstrated that the MUSE system can train huge merged models faster than individually built unmerged models on the main cloud. After a series of runs, we found that the combined model reduced training time by an average of 26.2%.

Introduction

Using interpretation, visualization, and statistical modeling, social network analysis (SNA) seeks to understand the underlying social structure of social networks. Patterns in a collection of connected items, such as people, can be studied and displayed using the many components that make up social data. SNA has emerged as a very active research topic, with many individuals and organizations devoting time and resources to it. According to many researchers¹⁻⁴, the SNA approach examines structures in attitudes, behaviour, and consequences among individuals, groups, and organisations. According to Hatala¹, the steps involved in implementing SNA are choosing the type of analysis, determining network relationships using theoretically plausible metrics, and collecting network data, measuring relationships, and attribute information that actors relate to, Analyze network data, create descriptive indexes, and present network data.



Bodin Ö, et al. 2020.
Annu. Rev. Environ. Resour. 45:471-95

Fig 1 Study of Social Networks and Healthcare Communications

However, the recent increase in corpus size and variety of domains and subdomains has made automatic classification difficult. Research areas that had only a limited field of view five years ago are now expanding rapidly and attracting strong interest. There are many fields, including biology (such as CRISPR-CA9), materials science (such as chemical programming), and health, are seeing this expansion in subfields (e.g. precision medicine). With the expansion of sub-subjects, it is important to organize documents by both the main subject and all related sub-subjects, rather than simply labeling them by subject. Hierarchical taxonomy is used here.

Critical service systems such as smart grids, next-generation healthcare, smart transportation, and utilities are gradually implementing IoT sensors and actuators, multi-cloud and virtual computing. network functionalization (NFV) [1], [2]. We call these next-generation systems to contrast them with older services. Advances in technology have obvious advantages, but they also expand the attack surface of targeted services, increasing the opportunities for adversary attackers to mount devastating attacks against these systems. According to a recent study, 75% of healthcare facilities worldwide face attacks [3]. US and UK cyber organizations issued warnings in May 2020 that hackers were taking a broad approach to sabotage research on the coronavirus (Covid-19) [4]. A survey of attacks on major utilities found that 56% of respondents reported having been attacked at least once in the last 12 months, resulting in data loss[5].

An organization's effectiveness in achieving specific goals is assessed using a maturity model. It may also help companies identify where their processes are good and where they are not. A tool called the Cybersecurity Maturity Model allows you to compare similar companies within your industry and track progress made over time as a result of integrating security into your organization's tactical and strategic work processes. increase.

Most current cybersecurity maturity models are based on assessments of compliance with cybersecurity policies and regulations, or estimates of network, vulnerability risk, and intrusion detection are examples of some (SI) components [4, 5]. And still, these evaluations approaches cannot accurately represent a healthcare organization's overall security posture evaluating the performance of individual IS components, and adherence to standards alone does not provide objective assessment results. considered ineffective. These discrepancies prevent its use as a model to generate reliable and measurable results [9]. Each healthcare institution's security posture creates transparency and builds trust among participating organizations so that interconnected healthcare systems can communicate effectively without compromising overall system security. should be well known using cyber security metrics.

RELATED WORK

Property abandonment in Buffalo, New York was predicted by Yin and Silverman (2015) using field data and GIS data. The abandonment and destruction of a heritage city in the United States was examined using cross-sectional models in this study.

Kumagai et al. (2016) used basic Bayesian utility data (fire hydrant data) to model housing abandonment. The link between closed faucets and vacant and abandoned houses has been confirmed. However, due to differences in data accessibility and quality between different cities, systems using topographic data are limited when applied to a wide geographic area. large (Deng and Ma, 2015). In two surveys, researchers used remote sensing data to determine the abandoned status of an individual dwelling by characterizing its physical characteristics (Deng and Ma, 2015; Zou and Ma, 2015; Zou et al. Wang, 2020).

According to Boukerche, Yen, Hwang and others, security and privacy are the most crucial issues in portable healthcare. Security is one of the most important concerns in mobile healthcare, and methods such as two-way telephony are being explored to address health metrics. Mobile healthcare includes telemedicine, data transmission, user mobility, and the use of wearable technologies (such as sensors and implants). move.

In a recent study, Li et al.⁴⁵ found that even if users are knowledgeable enough about privacy controls and can change their privacy settings, social media information can still be leaked. rust. According to Külcü and Henkolu⁴⁶, there is a need for greater social awareness of privacy protection and the lack of strong anti-cybercrime laws also helps to reduce opposition to privacy issues. It is clear that society needs mobile healthcare solutions, especially for older people who prefer to travel in an emergency.

Methodology Cloud security guidelines and best practises for the healthcare industry

Because they are reactive in nature, standards, guidelines, and the concept of best practises is not new. There will always be a lag between deciding something is necessary and actually putting it into practice, which can take years. Due to the different goals of many countries, this can widen the implementation gap , it becomes a bigger problem for international standards. The problem is exacerbated by the technological environment such as cybersecurity and especially by rapidly changing technology such as cloud technology. However, in addition to rapidly changing technology, the threat environment is also changing rapidly [10].

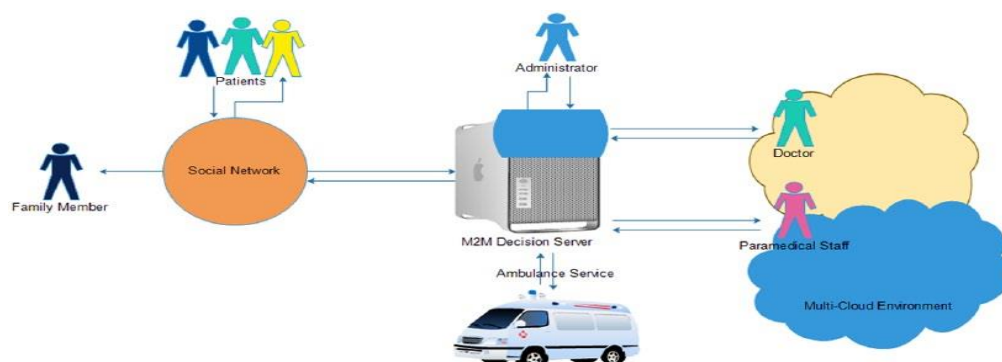


Figure 2. An overview of the medical system.

A summary of the system is presented in Figure 2, outlining potential interactions between system components to show the overall composition of the system.

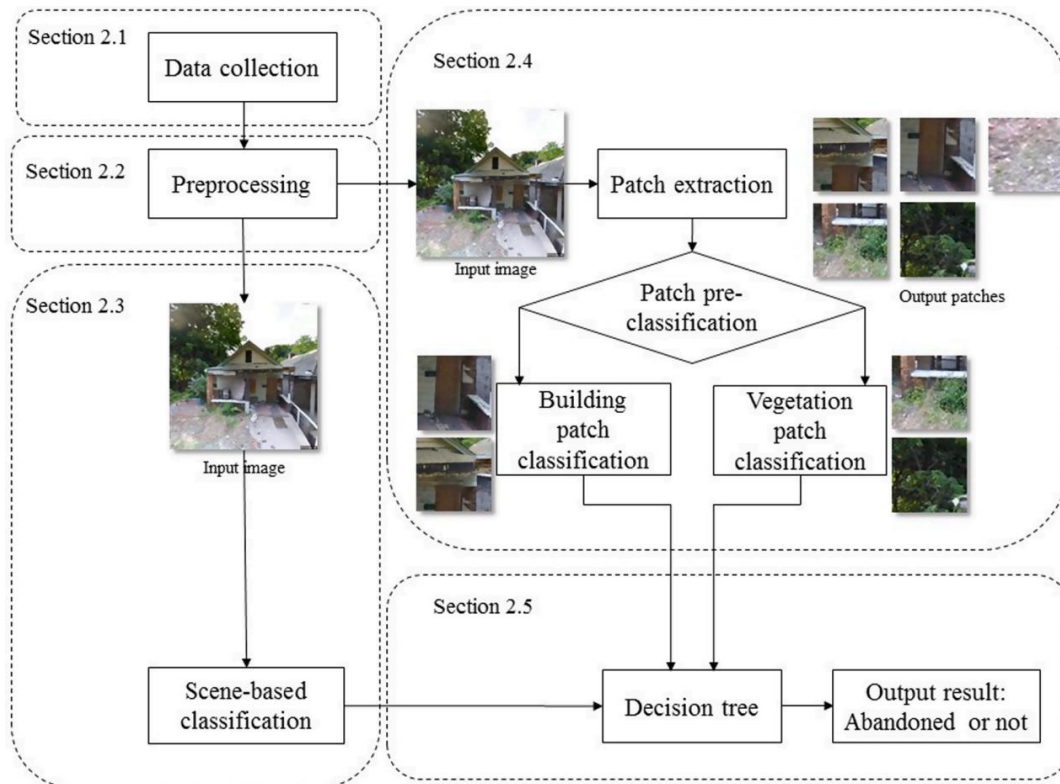


Fig. 3. AH detection by Street View images using hierarchical deep learning in workflow.

In this study, we developed a new deep learning hierarchical classification technique to distinguish AH images from GSV images (Figure 3). Global-level features and local-level features were extracted simultaneously using a scene-based classifier and a patch-based classifier, respectively, after data collection and pre-processing. A basic deep CNN model was trained as the basis for the scene-based classification procedure. Patches have been pre-generated and classified into three categories for patch-based classification: construction, vegetation and others. Construction land and vegetation were used to train two deep CNN models to separately identify degraded construction land and overgrown vegetation. The hierarchical classification of texts is the main contribution of this study. Traditional multi-class classification techniques can work well for a limited number of classes but will suffer as the number of classes increases, like with hierarchically classified documents. By creating designs that specialize deep learning techniques at each level of the document hierarchy, we address this in our hierarchical deep learning model (for example, see Figure 3.).

Hierarchical Deep Learning for Text (HDLTex) calculation

Our Hierarchical Deep Learning Architecture for Text (HDLTex) contains several deep learning models, each of which is organised as follows:

8 hidden layers in DNN, each with 1024 cells.

RNN: 100 GRU cells with two hidden layers are used in this implementation, with LSTM.

CNN: The CNN has 8 hidden layers and filter sizes of 3, 4, 5, 6, 7, a maximum pool of 5, layer sizes of 128, 128, and a maximum pooling of 5, 5, 35. The following variables were applied to all models: Batch size is 128 and the learning parameters are 0.001., $\theta_1=0.9$, $\theta_2=0.999$, $= 1e-08$, $decay = 0.0$, $Dropout=0.5$ (DNN) and $Dropout=0.25$ (CNN and RNN).

A. Evaluation

For the deep learning models, We employed the next cost function:

$$\text{Acc}(X) = \sum_{\varrho} \left[\frac{\text{Acc}(X_{\Psi_{\varrho}})}{\kappa_{\varrho}-1} \sum_{\Psi \in \{\Psi_1, \dots, \Psi_k\}} \text{Acc}(X_{\Psi_1}) \cdot n_{\Psi_k} \right] \quad (1)$$

where ϱ is the hierarchy's level count, k is the number of classes for each level, and Ψ is the number of classes in the child level.

B. Optimization

In this study, we applied two different kinds of stochastic gradient optimization for the deep learning models: RMSProp and Adam. Here is a description of these.

Optimizer for RMSProp: Below is a basic illustration of stochastic gradient descent (SGD):

$$\theta \leftarrow \theta - \alpha \nabla_{\theta} J(\theta, x_i, y_i) \quad (2)$$

$$\theta \leftarrow \theta - (\gamma \theta + \alpha \nabla_{\theta} J(\theta, x_i, y_i)) \quad (3)$$

In these equations, the learning parameter is ϑ , the learning rate is α , and the objective or cost function is $J(\vartheta, x_i, y_i)$. The definition of the update history is $\gamma \in (0, 1)$. The equation shows how SGD uses the impulse term with the rescaled gradient to update the parameters (3). The problem with sparse gradients is that this optimization method does not perform bias correction.

Adam Optimizer: As illustrated below, Adam is a different stochastic gradient optimizer that really only takes into account the gradient's first two moments, v and m .

$$\theta \leftarrow \theta - \frac{\alpha}{\sqrt{\hat{v}} + \epsilon} \hat{m} \quad (4)$$

Where

$$g_{i,t} = \nabla_{\theta} J(\theta_i, x_i, y_i) \quad (5)$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_{i,t} \quad (6)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_{i,t}^2 \quad (7)$$

The first and second moments in these equations are denoted by m_t and v_t , respectively.

Table 1. Compare this survey to the current composite framework for choosing cloud services.

Criteria	CloudGenius	CloudRecommender	Dastjerdi et.al	Proposed approach
Matching methods	AHP	SQL semantics	FCB8JGA + cost	SievingJGA + utility
Dynamic matching	No	No	No	Yes
Framework	CumulusGenius	CloudRecommender	CloudPick	CloudService Broker
Application scope	Multi-tier	Multi-tier	Multi-tier	Scientific workflows
SLA metrics/pricing	Synthetic	Real	Real	Real
Evaluation criteria	Time complexity	Cost + Time complexity	Cost + Time complexity	Cost + QOS + Time complexity
Ontology-based	No	Yes	Yes	Yes

Table 2 Multi-Cloud matching algorithm.

Parameters		Description
Request	$VM=(vm_1, \dots, vm_a)$	Requirements for various computing services
	$ST=(st_1, \dots, st_b)$	the set of b compute services requested
	$S=(s_1, \dots, s_n), S=VM \cup ST$	A collection of n component services forming a composite requirement
	$E=\{e_{s_i s_j} s_i, s_j \in S\}$	collection of all f edges

		connecting n services
	$L=\{lat(e_{s_i s_j}) s_i, s_j \in S\}$	A collection of edge latencies
	$Q_r = \{q_1(r), \dots, q_m(r)\}$	A group of m mandatory QoS values
	$Cr=\{C_{vm}(r), (C_{st}(r), C_{tr}(r)\}$	Maximum payment for computing power, storage and data traffic
	$D=\{d_{st}, d_{IT}\}$	Storage requirements and amount of data
	T	Deployment time (lease)
Provider	$P = \{p_1, \dots, p_l\}$	List of potential IaaS providers
	$Q_{pj} = \{q_1(p_j), \dots, q_m(p_j)\}$	Measured QoS metrics for provider p_j
	$C_{pj} = \{C_{vm}(p_j), C_{st}(p_j), C_{tr}(p_j)\}$	Pricing policies for provider p_j
Candidate	$X = \{x = \bigoplus s_1 \rightarrow p_1 s_i \in S, p_j \in P\}$	A collection of potential service configurations
	$Q_x = \{q_1(x), \dots, q_m(x)\}$	Service configuration QoS value set
	$C_x = \{c_{1m}(x), c_{st}(x), c_{tr}(x)\}$	total consumption cost of the composition x

Finding a service configuration that minimizes deployment costs and best meets the QoS requirements of all users is a multi-cloud matching problem. We presented a mathematical model for formulating this multi-objective NP-hard optimization problem. The latter serves as a model for configured Matching algorithms' input parameters include service demands, cloud providers, and potential service configurations.. Table 2 contains descriptions of the model parameters.

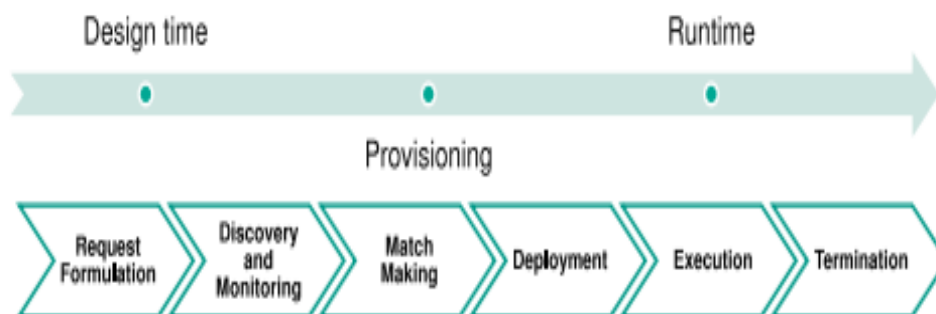


Fig. 4. Multi-Cloud brokering life cycle.

We made the following traffic assumptions for our application requirements in order to be able to transfer data across the necessary services during runtime:

Assumption 1:

The user gets paid only for her internet traffic from the cloud.

2. Assume data transfer within the same provider is free.

3. Assume that the total amount of data stored by all requested storage services is equal (data is replicated). 4. Suppose data is sent in both directions between two connected nodes.

5. Assume equal shares of connected clouds in the desired data flow.

Our mathematical approach allows us to formulate the matchmaking problem as follows:

$$\begin{cases} \max th(x), av(x) & \text{with } th(x) \geq th(r) \wedge av(x) \geq av(r) \\ \min rt(x), lat(x) & \text{with } rt(x) \leq rt(r) \wedge lat(x) \leq lat(r) \\ \min C_x(T) & \text{with } C_x(T) \leq C_x(T) \end{cases} \quad (8)$$

Where $Q_x = rt(x), th(x), av(x), lat(x)$ represent the set of minimum or maximum values permitted and $Q_r = rt(r), th(r), av(r), lat(r)$ represent the aggregated SLA values for response time, throughput, availability, and latency of composite service x . The former is determined as shown in Table 3 by using the aggregation functions given in [15] and the relevant values in the component services, where the data transfer charges are denoted by the symbol $C_x(T)$. With the above mentioned assumptions, $C_x(T)$ can be calculated as follows

$$C_x(T) = T * \left(\sum_{i=1}^a c_{vm}(vm_i) + d_{st} * \sum_{i=1}^b c_{st}(st_i) + \sum_{e \in E}^{s_k s_1 \in k \neq 1} \frac{d_{tr} * c_{tr}(e_{s_k s_1})}{2 * f} \right) \quad (9)$$

$$c_{tr}(e_{s_k s_l})_{s_k s_l \rightarrow p_m p_n} = \begin{cases} c_{tr}(p_m) + c_{tr}(p_n) & \text{if } m \neq n \\ 0 & \text{else} \end{cases} \quad (10)$$

where $Cr(T)$ means the total user budget for the usage period T and $p_m, p_n \in P$ denote the cloud provider s_k and $s_l \in S$ allocated for the requested service. The formula is:

where $p_m, p_n \in P$ denote the Cloud providers allocated respectively to the requested services $s_k, s_l \in S$. $Cr(T)$ denotes the total user budget for the period of usage T . It is calculated as follows:

$$C_r(T) = T * (a * c_{vm}(r) + b * d_{st} * c_{st}(r) + d_{tr} * c_{tr}(r)) \quad (11)$$

A weighted scoring algorithm is used to describe the user's preference for SLA features that do not work, but all functional requirements like sifting must be met. Use the following formula to calculate customer I's benefit from consuming potential multi-cloud federation service x over time T and ensuring quality of service Qx.

$$U_{ix}(Q_x, T) = C_{ri}(T) * F_i(Q_x) - C_x(T) \quad (12)$$

where $F_i(Q_x)$ is the customer merit function and $C_x(T)$ is the total cost of using the service (see Equation 9). $C_{ri}(T)$ reflects a customer I's greatest willingness to pay over a period T for "perfect" service quality. Here is the explanation.

$$F_i(Q_x) = \sum_{j=1}^m \lambda_i(q_j) * f_i(q_j) \rightarrow [0,1] \quad (13)$$

where $i(q_j)$ and $f_i(q_j)$ for $m_j=1$ and $i(q_j)=1$ represent the relative weighting and adjustment function of consumer I to SLA feature q_j respectively. Each evaluated SLA attribute maps to exactly one real value between [0, 1]. by the fitting function. where 1 indicates the ideal expected SLA value. It's a potential cloud service configuration, practical, and offers the best usage value.

$$U_{ix_{optimal}}(Q_x, T) = \max_{x \in X} U_{ix}(Q_x, T) \quad (14)$$

where X is a collection of potential configurations of cloud services (solution space). As a result, the matchmaking problem is to find the cloud configuration that gives the user the most benefit.

Table 3. Social media privacy and security issues.

Key Data Security Challenges in Social Networks	Top Privacy Concerns on Social Networks
1. Corporate impersonation	1. Privacy of user personal space
2. Customer scams	2. User identification
3. Traffic hijacking	3. User communication
4. Watering hole phishing and malware	
5. Account takeover	
6. Information leakage	
7. Click bait attacks	

The Social media privacy and security issues such as Key Data Security Challenges in Social Networks and Top Privacy Concerns on Social Networks and Validation Datasets 1 and 2 Performance of a hierarchical deep learning system in different confidence zones are shown in table 3 and 4.

Table 4.Validation Datasets 1 and 2 Performance of a hierarchical deep learning system in different confidence zones.

		Number of images				% (95% CI)	
Datasets	Regions	True Positive	False Negative	True Negative	False Positive	Sensitivity	Specificity
	Overall region	2770	114	3210	207	0.9605 (0.953-0.968)	0.9394 (0.931-0.947)
validation dataset 1	Reliable region	2019	48	2455	55	0.9768 (0.970-0.983)	0.9781 (0.972-0.984)
	Suspicious region	751	66	755	152	0.9192 (0.900-0.938)	0.8324 (0.808-0.857)
validation dataset 2	Overall region	592	27	1265	80	0.9564 (0.940-0.973)	0.9405 (0.928-0.953)
	Reliable region	493	8	929	17	0.9840 (0.973-0.995)	0.9820 (0.974-0.991)
	Suspicious region	99	19	336	63	0.8390 (0.772-	0.8421 (0.806-

						0.906)	0.878)
--	--	--	--	--	--	--------	--------

Result and Discussion

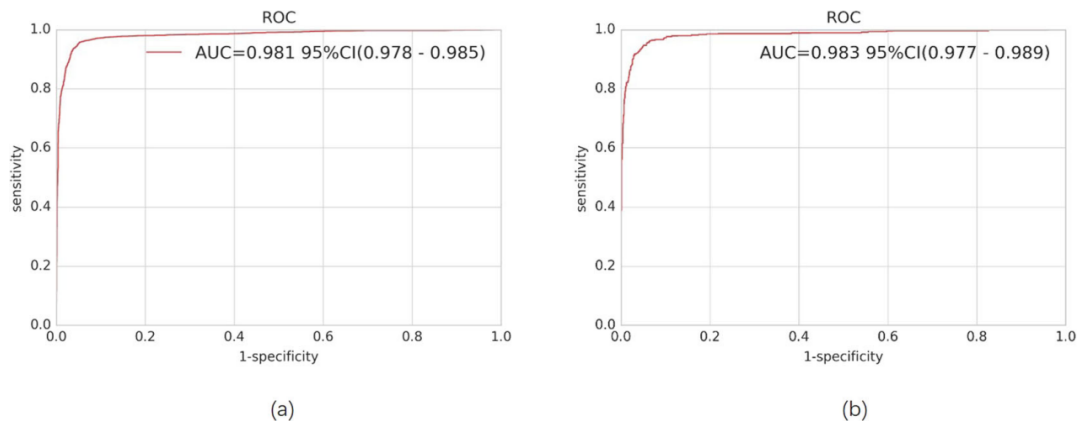
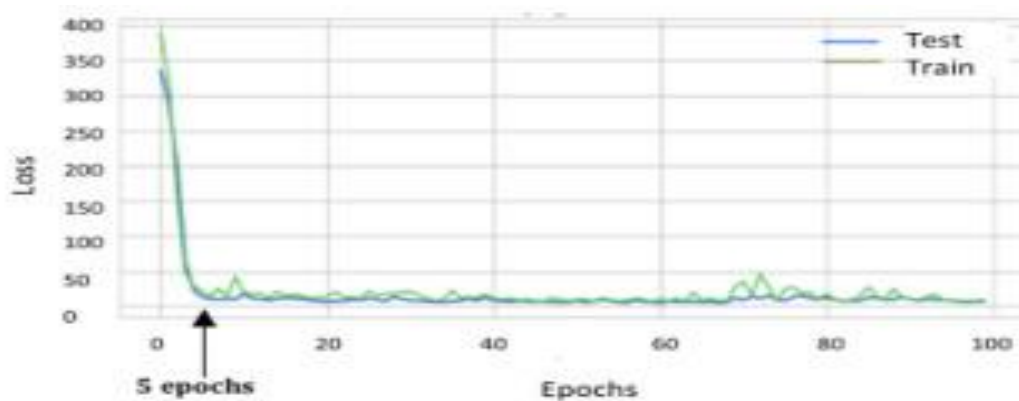


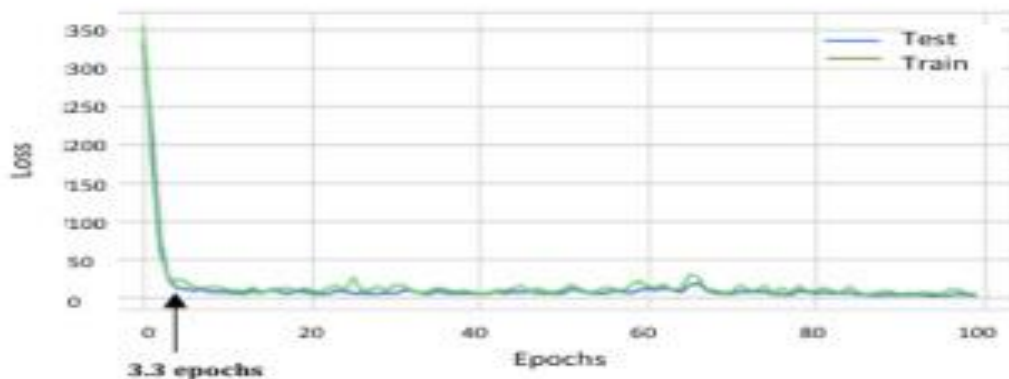
Figure .5 Graphs showing Receiver Operating Characteristic (ROC) curves derived from validation datasets 1(a) and 2 using the hierarchical deep learning system (b).

We evaluated the performance of the proposed HDLS using validation dataset 1. The receiver operating characteristic (ROC) curve is shown in Figure 5. Area under the ROC curve (AUC) is 0.981 (95% CI, 0.978-0.985), as can be seen. Approaches sensitivity and specificity were 96.1% (95% CI, 95.3-96.8%) and 93.9% (95% CI, 93.1-94.7%, respectively). , when the threshold is set at 0.8, as shown in Table 4. Evaluation of the effectiveness of the corresponding HDLS method in the reliable and problematic regions. Compared with results in the doubtful region, the sensitivity and specificity in the confidence region can reach 97.7% (95% CI, 97.0-98.3%) and 97.8% (95% CI, 97.2-98.4%). However, the sensitivity and specificity could still be 91.9% (95% CI, 90.0-93.8%) and 83.2% (95% CI, 80.8-85.7%) in the suspect area.



Train-loss = 4.517, Test-loss = 6.151 Train-accuracy = 0.989, Test-accuracy = 0.990

(a) Core cloud merged model

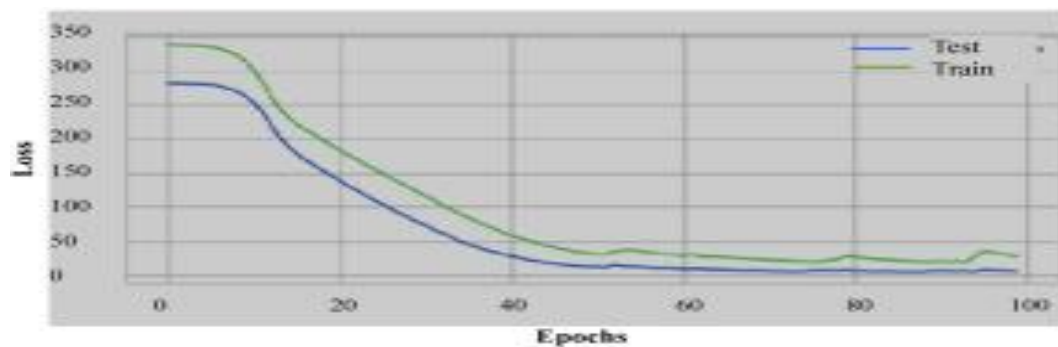


Train-loss = 2.1575, Test-loss = 4.1718,
Train-accuracy = 0.992, Test-accuracy = 0.993

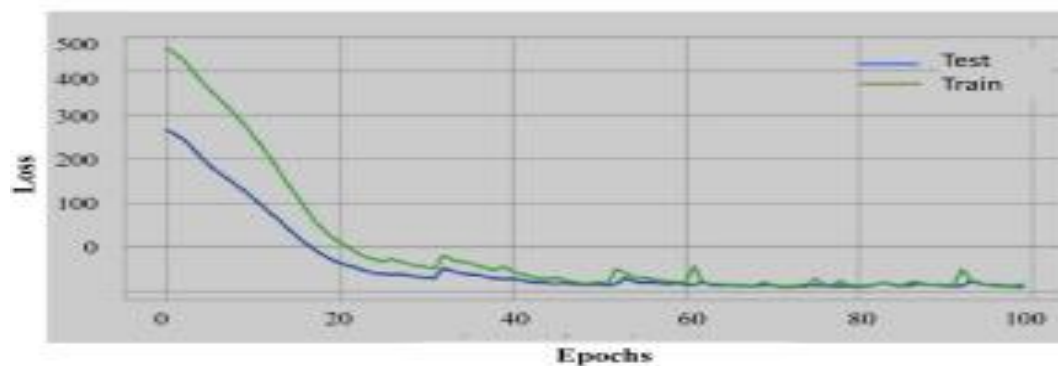
(b) Core cloud merged model (cross-trained)

Figure 6. A hierarchical deep learning approach was used to test the train, test, and train accuracy as well as test accuracy of edge clouds 1, 2, and 3 of a multi-cloud healthcare system.

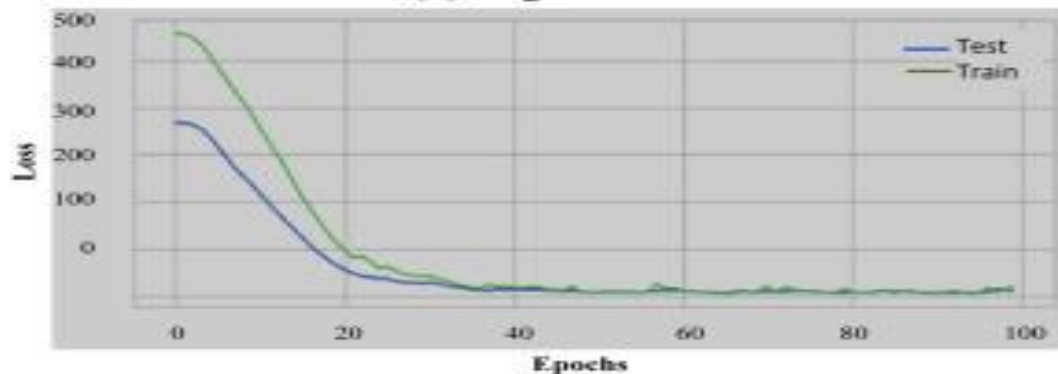
In addition, the researchers found that to complete the process within the two-day rental period, more than 30 virtual machines were required; If this requirement is not met, the user will be charged for the additional time required. The actual makespan differs from the theoretical one due to the overhead created by the file transfer, especially when multiple virtual machines are involved in the aggregation service. Figures 6 and 7 illustrate how edge clouds 1, 2, and 3 of a multi-cloud medical system are tested for instruction loss, test loss, training accuracy, and test accuracy when using hierarchical deep learning.



(a) Edge cloud 1



(b) Edge cloud 2



(c) Edge cloud 3

Fig. 7 Augmentation In a hierarchical deep learning technique, edge clouds 1, 2, and 3 of a multi-cloud medical system show training loss, test loss, training accuracy, and test accuracy. check.

Conclusion

This work aims to provide a technique that can be used to extract individual-level AH from GSV data. We have developed a completely new hierarchical deep learning technique that can detect AH based on global and local visual features. Data collection from IoT devices, analysis and storage across multiple clouds, and virtualization of communication components are on the rise. As technology advances, these systems become more vulnerable to hacker attacks. The purpose of the research described in this article is to demonstrate the feasibility and use of hierarchical, sparse deep neural network models based on artificial intelligence for data-in-motion security. The result is faster training times and more accurate detection of unconfirmed attacks. The core cloud model converges in much less time (6-8) than the boundary cloud model, which takes 35-40 epochs to train, and uses much less data. A comprehensive evaluation shows that the MUSE system can train large merged models faster than independently built unmerged models on the main cloud.

Reference

- [1] Jafari S, Mtenzi F, Fitzpatrick R, O'shea B (2010) Security Metrics for e-Healthcare Information Systems: A Domain Specific Metrics Approach. *Int J Digit Soc* 1:
- [2] Herrmann DS (2007) Complete Guide to Security and Privacy Metrics: Measuring Regulatory Compliance, Operational Resilience, and ROI, 1st ed. Auerbach Publications, London
- [3] Payne SC (2007) A Guide to Security Metrics
- [4] Pamula J, Ammann P, Jajodia S, Ritchey R (2006) A framework for establishing, assessing, and managing trust in inter-organizational relationships. In: *Proceedings of the 3rd ACM workshop on Secure web services - SWS '06*. ACM Press, New York, New York, USA, p 23
- [5] Chew E, Swanson M, Stine K, et al (2008) Performance Measurement Guide for Information Security: NIST Special Publication 800-55 Revision 1. Gaithersburg, MD
- [6] D. Chen, Z. Liu, L. Wang, M. Dou, J. Chen, H. Li, Natural disaster monitoring with wireless sensor networks: a case study of data-intensive applications upon low-cost scalable systems, *Mobile Netw. Appl.* 18 (2013) 651–663.
- [7] G. Juve, E. Deelman, Scientific workflows in the cloud, in: *Computer Communications and Networks*, Springer, London, 2011.
- [8] Y. Ma, L. Wang, A.Y. Zomaya, D. Chen, R. Ranjan, Task-tree based large-scale mosaicking for remote sensed imageries with dynamic DAG scheduling, *IEEE Trans. Parallel Distrib. Syst.* (2013).
- [9] M.C. Schatz, B. Langmead, S.L. Salzberg, Cloud computing and the DNA data race, *Nature Biotechnol.* 28 (2010) 691.

- [10] F.F. Costa, Big data in biomedicine, *Drug Discov. Today* (2013).
- [11] Shahzad B, Alwagait E and Alim S. Impact of change in weekend days on social networking culture in Saudi Arabia. In: *2014 international conference on future internet of things and cloud (FiCloud)*, Barcelona, 27–29 August 2014, pp. 553–558. New York: IEEE.
- [12] Shahzad B and Alwagait E. Smartphone’s popularity measurement by investigating Twitter profiles. In: *2013 third world congress on information and communication technologies (WICT)*, Le Quy Don University, Hanoi, Vietnam, 15–18 December 2013, pp. 349–352. New York: IEEE.
- [13] Al-Ohali Y, Al-Oraij AA and Shahzad B. KSU news portal: a case study. In: *The 2011 international conference on internet computing (ICOMP’11)*, Las Vegas, Nevada, USA, 18–21 July 2011, pp. 1–5. Georgia, USA: CSREA Press.
- [14] Griffiths F, Cave J, Boardman F, et al. Social networks—the future for health care delivery. *Soc Sci Med* 2012; 75: 2233–2241.
- [15] Sarasohn-Kahn J. *The wisdom of patients: health care meets online social media*. Oakland, CA: California HealthCare Foundation, 2008, pp. 1–27.
- [16] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [17] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, “Learning and transferring mid-level image representations using convolutional neural networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1717–1724.
- [18] J. Y. Lee and F. Dernoncourt, “Sequential short-text classification with recurrent and convolutional neural networks,” *arXiv preprint arXiv:1603.03827*, 2016.
- [19] X. Zhang, J. Zhao, and Y. LeCun, “Character-level convolutional networks for text classification,” in *Advances in neural information processing systems*, 2015, pp. 649–657.
- [20] L. Medsker and L. Jain, “Recurrent neural networks,” *Design and Applications*, vol. 5, 2001.
- [21] R. Salakhutdinov, J. B. Tenenbaum, and A. Torralba, “Learning with hierarchical-deep models,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1958–1971, 2013.
- [22] Fu, H. et al. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. *IEEE Trans. Med Imaging* 37, 1597–1605 (2018).
- [23] Sammut, C. & Webb, G. I. *Encyclopedia of Machine Learning*, (Springer US, 2010).
- [24] Qi, Z., Wang, B., Tian, Y. & Zhang, P. When ensemble learning meets deep learning: A new deep support vector machine for classification. *Knowl.-Based Syst.* 107, 54–60 (2016).

- [25] Angermueller, C., Pärnamaa, T., Parts, L. & Stegle, O. Deep learning for computational biology. *Mol. Syst. Biol.* 12, 878 (2016).
- [26] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. in 2016 IEEE Conference on Computer Vision and Pattern Recognition (2016).