

A Vertical Handoff Algorithm Based on Mobile User Location Prediction (Postprint)

Authors: Meng Lei, Tang Xin, Xu Yanyan

Date: 2022-04-07T15:01:56Z

Abstract

To address the issue that existing vertical handover algorithms in heterogeneous wireless network environments struggle to achieve seamless handover for mobile users accessing networks and fail to provide stable communication services, a vertical handover algorithm based on mobile user location prediction is proposed. First, by leveraging the similarity of user mobility trajectories, an LSTM model is offline trained using user trajectory data to learn the universal mobility patterns inherent to various users. Subsequently, the LSTM model is loaded online to perform user location prediction, thereby enabling the calculation of reward values for candidate networks corresponding to users at the next time instant through fuzzy logic analysis, and selecting the network with the highest reward value for handover. Experimental results demonstrate that, compared with existing vertical handover algorithms, the proposed algorithm achieves significant reductions in the number of handovers, handover failures, and handover latency under varying user mobility speeds and user scales, while maintaining low memory consumption, thus enabling seamless handover.

Full Text

Preamble

Vol. 39 No. 8

Application Research of Computers

Accepted Paper

A Vertical Handoff Algorithm Based on Mobile User Location Prediction

Meng Lei, Tang Xin, Xu Yanyan[†]

(State Key Laboratory of Information Engineering in Surveying, Mapping & Remote Sensing, Wuhan University, Wuhan 430079, China)

Abstract: Existing vertical handoff algorithms in heterogeneous wireless network environments struggle to achieve seamless handoff for mobile users and fail to provide stable communication services. To address these issues, this paper proposes a vertical handoff algorithm based on mobile user location prediction. First, leveraging the similarity in user mobility trajectories, the algorithm offline trains an LSTM model using user trajectory data to learn the general movement patterns shared across users. It then online loads the LSTM model to predict user locations, employs fuzzy logic to calculate the reward values of candidate networks for the next time instant, and selects the network with the highest reward for handoff. Experimental results demonstrate that compared with existing vertical handoff algorithms, the proposed algorithm significantly reduces the number of handoffs, handoff failures, and handoff delay under varying user mobility speeds and scales, while maintaining low memory consumption, thereby enabling seamless handoff.

Keywords: vertical switching; LSTM; fuzzy logic; heterogeneous network

0 Introduction

Mobile users moving between heterogeneous wireless networks such as cellular mobile communication networks and WLAN (Wireless Local Area Network) coverage areas frequently switch networks, making it difficult to guarantee communication stability. Vertical handoff refers to switching between dissimilar current and target access networks. Efficient vertical handoff algorithms are essential prerequisites for achieving seamless handoff for mobile users in heterogeneous wireless networks and serve as critical enablers for providing stable communication services.

SINR-based vertical handoff algorithms select the network with the highest Signal-to-Interference-plus-Noise Ratio (SINR) value for switching, but may cause frequent handoffs among multiple networks with similar SINR values, degrading quality of service. Multi-attribute decision theory-based vertical handoff algorithms abstract the handoff process as a multi-attribute decision-making problem, employing fuzzy logic for comprehensive evaluation based on network attribute parameters to avoid frequent handoffs. However, these methods struggle to provide accurate location predictions during user movement, cannot reserve handoff resources in advance, and may lead to handoff failures due to insufficient resources, thereby increasing unnecessary handoffs and offering limited performance improvement. Markov model-based location prediction methods can provide user location information to assist handoff, but in heterogeneous wireless networks, their state space scales grow dramatically with the number of networks, causing computational complexity to surge and increasing prediction delay, which degrades handoff performance. Personalized location prediction-based vertical handoff algorithms establish Long Short-Term Memory (LSTM) models for each user to learn individual mobility patterns and perform person-

alized location prediction to assist handoff, avoiding the exponential growth in computational complexity with network count. However, memory and computational costs increase linearly with the number of users, making them impractical for multi-user scenarios. In summary, existing vertical handoff algorithms in heterogeneous wireless networks face challenges in achieving seamless handoff and providing stable communication services.

Therefore, this paper proposes a vertical handoff algorithm based on mobile user location prediction. By exploiting the similarity in user mobility trajectories, the algorithm transforms personalized location prediction into a unified prediction model for all users, thereby reducing computational costs. Second, the historical mobility trajectories of mobile users undergo grid-based preprocessing to reduce LSTM prediction difficulty and improve accuracy. Subsequently, the trained LSTM model predicts users' next locations online, and fuzzy logic analyzes the corresponding network attribute parameters at these predicted locations to select the optimal future network for handoff, ensuring stable network services.

1 Model Framework

The proposed vertical handoff algorithm adopts a centralized controller design inspired by SDN (Software Defined Network) architecture, deploying the LSTM model and fuzzy logic controller within the centralized controller. This addresses the computational requirements of the LSTM model while enabling the fuzzy logic controller to fully utilize global network attribute information for handoff decisions. During offline training, user mobility trajectory information is collected from the simulated network, preprocessed via gridding, and used as a dataset to train the LSTM model.

Consider the heterogeneous network scenario illustrated in Figure 1 to describe the online model framework for handling network handoffs. The 4G network provides wide-area coverage, while WLAN networks offer high data rate services within small areas. Users move through the heterogeneous network, continuously switching to new networks. The centralized controller loads the LSTM model trained offline using collected historical user mobility data. The online handoff process proceeds as follows:

- a) **Information Collection:** The centralized controller periodically collects user location and mobility speed information for location prediction.
- b) **Location Prediction:** The controller combines real-time user information with historical trajectory data, preprocesses it through gridding, and feeds it into the trained LSTM model to predict the user's location at the next time instant.
- c) **Network Parameter Estimation:** Using the predicted location, the controller estimates multi-attribute parameters (e.g., received signal strength, Doppler shift) of candidate networks at the next time instant.

- d) **Multi-Attribute Parameter Processing:** Fuzzy logic processes the estimated network multi-attribute parameters to obtain future rewards for each candidate network.
- e) **Future Reward Analysis:** The centralized controller selects the network with the highest future reward for handoff decision.
- f) **Decision Update:** Mobile users execute the updated handoff decision from the centralized controller.

Figure 1 The scene diagram and model framework of handoff

2.1 Mobility Model

To ensure that user mobility trajectories in the simulated network realistically reflect human movement patterns, the SLAW (Self-Similar Least-Action Walk) mobility model is employed. The centralized controller's sampling period is T . Let $P_i(t)$ denote the position of user i at time t . The user's movement process is as follows:

- a) **Initialization:** Randomly select waypoints from the movable area to generate a candidate waypoint set W_i , then randomly choose two waypoints from W_i as the user's starting point and first destination, setting the starting point as the current waypoint.
- b) **Visiting Destination:** User i moves toward the destination at speed v_i .
- c) **Determining Dwell Time:** Upon reaching a destination, the user's dwell time follows a truncated power-law distribution.
- d) **Selecting Next Destination:** Since users prefer shorter paths, the probability of moving from the current waypoint $w_{i,c}$ to other candidate waypoints is calculated based on distance weighting. For user i currently at waypoint $w_{i,c}$, the probability of selecting candidate waypoint $w_{i,k}$ as the next destination is given by Equation (1):

$$P_i(w_{i,c}, w_{i,k}) = \frac{d_{i,c,k}^{-\beta}}{\sum_{k' \in W_i} d_{i,c,k'}^{-\beta}}$$

where W_i is the set of all candidate waypoints, $d_{i,c,k}$ is the Euclidean distance between current current waypoint $w_{i,c}$ and candidate waypoint $w_{i,k}$, and β is the distance weighting factor. A larger β indicates stronger preference for nearby waypoints; when $\beta = 0$, the user randomly selects the next destination from W_i . The waypoint with highest probability is chosen as user i 's next destination.

- e) **Loop:** Repeat steps b)-d) until the total movement time for user i is reached.

2.2 Location Prediction

The centralized controller collects user location information with period T . However, raw location data points are relatively dense, making it difficult to extract mobility patterns and degrading vertical handoff performance. Therefore, a commonly used gridding method for mobility trajectory research is adopted to preprocess raw trajectory data, converting precise location points into more abstract grid information to facilitate pattern mining.

As shown in Figure 2, solid and dashed lines represent raw trajectories L_1 and L_2 of two users. While the raw trajectories appear unrelated, gridding reveals both follow the pattern $g_1 \rightarrow g_2 \rightarrow g_3 \rightarrow g_4$, where g_i denotes a grid with index i . This converts precise location information into grid indices.

Figure 2 The gridding process of trajectories

Since gridded trajectories help analyze user mobility patterns and network attributes do not vary significantly within small grids, appropriate grid partitioning does not affect comprehensive network evaluation.

Let the network coverage area be an $X \times Y$ rectangle with grid width d_w . The user's position $P_i(t)$ at time t is mapped to grid index $g_i(t)$, yielding the grid index set $G_i = \{g_i(1), g_i(2), \dots\}$.

LSTM preserves historical location information in memory cells and uses gating mechanisms to control information retention. Let $g_i(t)$ be the input grid position of user i at time t , $f_i(t)$ the forget gate output, $I_i(t)$ the input gate output, $\tilde{C}_i(t)$ the candidate memory cell state, $C_i(t)$ the memory cell state, and $O_i(t)$ the output gate output. The LSTM-predicted next grid position $\hat{g}_i(t+1)$ is computed via Equation (2):

$$\begin{aligned} f_i(t) &= \sigma(W_f \cdot [h_i(t-1), g_i(t)] + b_f) \\ I_i(t) &= \sigma(W_i \cdot [h_i(t-1), g_i(t)] + b_i) \\ \tilde{C}_i(t) &= \tanh(W_c \cdot [h_i(t-1), g_i(t)] + b_c) \\ C_i(t) &= f_i(t) \cdot C_i(t-1) + I_i(t) \cdot \tilde{C}_i(t) \\ O_i(t) &= \sigma(W_o \cdot [h_i(t-1), g_i(t)] + b_o) \\ h_i(t) &= O_i(t) \cdot \tanh(C_i(t)) \end{aligned}$$

where W_f, W_i, W_c, W_o are weight matrices, b_f, b_i, b_c, b_o are bias terms, $h_i(t-1)$ is the output at time $t-1$, and σ is the activation function.

2.3 Network Model

Consider a heterogeneous wireless network environment with M available networks, each characterized by N multi-attribute parameters. This algorithm employs parameters including received signal strength (RSS) R_{ssi} , Doppler shift f_{dop} , signal-to-interference-plus-noise ratio (SINR) S_{nri} , channel capacity C_{max} ,

network delay d_{net} , and delay jitter j_{net} to characterize networks. The centralized controller collects user location $P_i(t)$ and speed $v_i(t)$ with period T , uses LSTM to predict user i 's location at time $t+1$ based on time t , and estimates the corresponding network parameters $R_{ssi}(t+1)$, $f_{dop}(t+1)$, $S_{nri}(t+1)$, $C_{max}(t+1)$ for fuzzy logic handoff decisions. Since delay and jitter cannot be estimated from location, fixed parameters are used for these metrics.

2.4 Fuzzy Logic Handoff Decision

In heterogeneous networks, different networks have varying parameter ranges, making direct numerical comparison impractical. Therefore, a fuzzy logic control model analyzes the multi-attribute parameters of heterogeneous networks to produce a unified evaluation output for objective candidate network assessment.

Let n be any network in user i 's candidate network set. The fuzzy logic analysis for comprehensive evaluation of network n 's multi-attribute parameters at time t proceeds as follows:

- 1) **Normalization:** Equation (3) normalizes the benefit of any attribute m for network n :

$$\text{benefit}_{m,n,i}(t+1) = r_{m,n} \times \frac{q_{m,n,i}(t+1) - \min(Q_{m,n,i})}{\max(Q_{m,n,i}) - \min(Q_{m,n,i})}$$

where $Q_{m,n,i}$ is the set of m -th attribute values for candidate network n of user i , $q_{m,n,i}(t+1)$ is the estimated m -th attribute value of network n at time $t+1$, and $r_{m,n}$ is the network benefit adjustment factor (negative for parameters where larger values indicate worse state, such as delay; positive for parameters where larger values indicate better state, such as RSS).

- 2) **Fuzzy Decision Mapping:** To balance computational overhead and decision effectiveness, triangular membership functions partition network parameters into three fuzzy sets: {Weak, Medium, Strong}. $z_{m,n}$ represents the fuzzy set for the m -th attribute value of network n .
- 3) **Network Parameter Benefit Weighting:** Since different parameters impact network state differently, the Analytic Hierarchy Process (AHP) constructs a decision factor matrix. The impact hierarchy of network state changes due to user mobility is set as: Doppler shift > SINR > RSS. Using the 1-9 scale method yields the decision factor matrix U in Equation (4). Consistency checking and adjustment ensure reliability, avoiding subjective bias.

$$U = \begin{pmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,N} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{N,1} & u_{N,2} & \cdots & u_{N,N} \end{pmatrix}$$

where N is the number of network parameter types. The normalized weight vector $w = \{w_1, w_2, \dots, w_N\}$ is derived from U .

- 4) **Network Reward Calculation:** Using the computed benefit weights, Equation (5) calculates the total multi-attribute reward $T_{benefit,i}(t+1)$ for network n at predicted time $t+1$ for user i :

$$T_{benefit,i}(t+1) = \sum_{m=1}^N w_m \cdot \text{benefit}_{m,n,i}(t+1)$$

The centralized controller obtains the rewards for all candidate networks at time $t+1$ and selects the network with highest reward n' for handoff execution.

2.5 Handoff Process

To effectively learn and utilize user mobility patterns, historical mobility dataset D is collected to offline train the LSTM model. Online, the centralized controller loads the trained model, uses real-time trajectory changes of user i to predict the next location, analyzes optional network rewards, and executes handoff. The detailed process is shown in Algorithm 1.

Algorithm 1 Handoff Decision Based on Mobile User Location Prediction

Input: Historical mobility dataset D , M candidate networks, N multi-attribute parameters, grid width d_w , sampling period T , user speed v_i

Output: Handoff to network n'

- a) Initialize grid width d_w and sampling period T .
- b) Grid-process dataset D and train LSTM model.
- c) Centralized controller loads the trained LSTM model online.
- d) **while** true **do**
 - e) Collect real-time trajectory $P_i(t)$ of user i , grid-process to obtain $g_i(t)$.
 - f) Input $g_i(t)$ into LSTM model to predict user i 's location at $t+1$: $\hat{P}_i(t+1)$.
 - g) Estimate candidate network parameters and compute rewards $T_{benefit,i}(t+1)$.
 - h) Select network n' with highest reward from candidate set and execute handoff.
- e) **end while**

One online handoff cycle comprises steps f-h. The location prediction complexity (step f) does not increase with the number of networks M or users. Steps g

and h have complexity $O(M)$, yielding overall handoff complexity $O(M)$, which satisfies real-time requirements.

3 Experiments

3.1.2 Training Process

The experimental platform uses an Intel Core i7-10710U CPU, Ubuntu 16.04 OS, and 32GB RAM. The centralized controller loads the trained LSTM model to asynchronously process location information for handoff decisions. To facilitate evaluation, the simulation simplifies the network environment to widely-used 4G and WLAN networks. The area size is 500×500 m, with varying numbers of mobile users and speeds while other parameters remain fixed. Initially, each user's position is randomly generated and connected to the network with highest RSS.

Table 1 Network Simulation Parameters

Parameter	4G Cellular Base Station	WLAN Network
Bandwidth	10MHz~30MHz	25MHz~40MHz
Delay	10ms~50ms	2ms~10ms
Transmit Power	23dBm	10dBm
Delay Jitter	40ms~100ms	15ms~30ms
RSS Threshold	-65dBm	-80dBm

Four metrics evaluate algorithm performance: handoff delay, memory consumption, handoff count, and handoff failure count. Lower handoff delay better meets seamless handoff requirements. Stable memory consumption with increasing users indicates practical applicability. Fewer handoffs reduce delay and improve service quality during switching. Fewer handoff failures indicate sufficient network capacity or strong signal strength, ensuring stable service quality.

The SLAW model generates user mobility trajectories with varying user counts and speeds. The training dataset contains 10,000 samples; the test set contains 2,000. LSTM weights are initialized using the Xavier method, biases set to zero. The network has three hidden layers: tanh activation for layers 1-2, ReLU for layer 3. Training uses RMSProp optimizer with gradient descent, 100 epochs, batch size 64, and learning rate 0.05. With $d_w = 3$ m, training error decreased from 1.08 to 0.007, achieving 94.1% test accuracy.

3.1.3 Grid Parameter Selection

The algorithm first grids user trajectories. Narrower grid width increases sensitivity to location changes and prediction accuracy, but network attribute variation becomes less distinct, limiting performance gains while increasing computational overhead. Since reducing grid width increases management difficulty

and cost, and network attributes barely change within 3m, the minimum d_w is set to 3m. Experiments compare $d_w = 9, 6, 3$ m.

Figure 3 shows handoff failure counts versus user count for different grid widths at speeds 1-5m/s. Smaller grid width yields fewer failures due to higher prediction accuracy, more precise future network parameter estimation, and more efficient fuzzy logic decisions. Balancing computational cost and accuracy, $d_w = 3$ m is selected.

3.2.1 Handoff Delay and Memory Consumption Comparison

To verify seamless handoff capability for multiple users, Figure 4 compares handoff delay across algorithms for varying user counts. All algorithms show stable delay with increasing users, but the proposed algorithm reduces delay by 48.36%, 33.53%, 47.12%, and 45.92% on average compared to VHO-SINR, UDWS, MPVH, and personalized prediction algorithms, respectively. The centralized control strategy enables asynchronous location prediction and handoff decisions using a single LSTM model, keeping delay low and user-count independent. VHO-SINR and UDWS cannot predict future locations, leading to suboptimal network selection and higher delay. MPVH's high computational complexity increases prediction delay. Personalized prediction requires per-user model loading, increasing delay.

Memory consumption comparison appears in Figure 5. Personalized prediction, MPVH, and UDWS memory usage grows with user count, while VHO-SINR and the proposed algorithm remain stable. The proposed algorithm maintains low memory usage, averaging 23.40%, 35.64%, 49.59%, and 61.89% reduction compared to personalized prediction, VHO-SINR, UDWS, and MPVH, respectively. This is because the proposed algorithm loads only one LSTM model for all users, whereas personalized prediction loads per-user models, MPVH's state space scales with users, and UDWS stores increasing historical attribute information, making them impractical for multi-user scenarios.

In summary, the proposed algorithm maintains low handoff delay and computational cost as user scale increases, meeting practical application requirements.

3.2.2 Handoff Performance Under Different User Speeds

To evaluate performance across mobility speeds, Figures 6 and 7 compare handoff counts and failures for 200 users at network density 75/km². The proposed algorithm reduces handoff counts by 43.09%, 19.15%, and 30.78% and handoff failures by 42.22%, 80.84%, and 26.11% on average compared to VHO-SINR, UDWS, and MPVH, respectively. VHO-SINR and UDWS cannot predict future behavior, triggering frequent handoffs based on current location. MPVH's Markov model considers only the previous location, yielding lower prediction accuracy than LSTM and higher failure rates. The LSTM-based approach accurately predicts next locations across speeds, reducing both handoff count and failures.

3.2.3 Handoff Performance Under Different User Counts

Figures 8 and 9 evaluate performance for varying user counts at network density $75/\text{km}^2$ and speed 1-5m/s. As user count increases, all schemes show rising handoff counts and failures. The proposed algorithm reduces handoff counts by 49.73%, 35.51%, and 33.84% and failures by 33.35%, 82.38%, and 12.3% compared to VHO-SINR, UDWS, and MPVH, respectively. VHO-SINR's preference for highest SINR causes network overload as users increase. UDWS cannot predict future locations, leading to suboptimal selection and more failures. MPVH's single-RSS criterion may concentrate users on popular networks, causing overload. The proposed algorithm's LSTM-based accurate prediction combined with fuzzy logic analysis of multi-attribute parameters prevents over-reliance on single metrics and avoids overloading popular networks through global reward maximization, thus reducing handoffs and failures.

4 Conclusion

To address the challenges of achieving seamless handoff and stable communication services in heterogeneous wireless networks, this paper proposes a vertical handoff algorithm based on mobile user location prediction. The approach transforms personalized prediction into a unified model for all users to reduce computational cost, employs grid-based preprocessing of historical trajectories to improve LSTM prediction accuracy, and uses the trained LSTM model for online location prediction. Fuzzy logic analyzes network attribute parameters at predicted locations to select the optimal future network, ensuring stable services. Experimental results demonstrate significant reductions in handoff count, failures, and delay across varying user speeds and scales, with low memory consumption, enabling seamless handoff.

References

- [1] Liu Jiajia, Shi Yongpeng, Fadlullah Z M, et al. Space-Air-Ground Integrated Network: A Survey [J]. IEEE Communications Surveys Tutorials, 2018, 20 (4): 2714-2741.
- [2] Li Sisi, Cheng Lianglun, Wang Tao, et al. Vertical handoff decision algorithm for multi-terminal cooperative in mobile interconnected heterogeneous networks [J]. Application Research of Computers, 2018, 35 (08): 2474-2476+2497.
- [3] Ayyappan K, Narasimman K, Dananjayan P. SINR Based Vertical Handoff Scheme for QoS in Heterogeneous Wireless Networks [C]// 2009 International Conference on Future Computer and Communication. Piscataway, NJ: IEEE Press, 2009: 117-121.
- [4] Yu Hewei, Ma Yanan, Yu Jingxi. Network Selection Algorithm for Multiservice Multimode Terminals in Heterogeneous Wireless Networks [J]. IEEE Access, 2019, 7: 46240-46260.
- [5] Goyal P, Lobiyal D K, Katti C P. Dynamic user preference based group vertical handoffs in heterogeneous wireless networks: a non-cooperative game

- approach [J]. *Wireless Networks*, 2020, 26 (2): 775–789.
- [6] Wang Jibo, Yang Lei, Qi Dongyuan, et al. The Vertical Handoff Research Based on Forward Differential Prediction and Fuzzy Logic [J]. *Application Research of Computers*, 2020, 37 (S2): 284-286.
- [7] Zhao Hu, Eyhab A. A User-Centered Handoff Procedure in 5G Cellular Networks [C]// *Companion Proceedings of the Web Conference 2020*. Association for Computing Machinery. New York: ACM Press, 2020: 1-5.
- [8] Mahmoud A, Ghaith H, Wail M. Improving Vertical Handoffs Using Human Mobility Prediction [J]. *International Journal of Advanced Computer Science and Applications*, 2016, 7 (3): 413-419.
- [9] Wang Chujie. Research on Mobility Prediction Methods and Resource Management Based on User Mobility [D]. Hangzhou: ZheJiang University, 2020.
- [10] Althé F, de La Fortelle A. An LSTM network for highway trajectory prediction [C]// *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. Piscataway, NJ: IEEE Press, 2017: 353-358.
- [11] Zhang Jin. The Management Research Based on User Mobility Prediction in Heterogeneous Wireless Networks [D]. Beijing: Beijing University of Posts and Telecommunications, 2015.
- [12] Wang Chujie, Ma Lin, Li Rongpeng, et al. Exploring Trajectory Prediction Through Machine Learning Methods [J]. *IEEE Access*, 2019, 99: 101441-101452.
- [13] Wang Chujie, Zhao Zhifeng, Qi Sun, et al. Deep Learning-based Intelligent Dual Connectivity for Mobility Management in Dense Network [C]// *IEEE 88th Vehicular Technology Conference (VTC-Fall)*. Piscataway, NJ: IEEE Press, 2018: 1-5.
- [14] Yu Yinbo, Li Xing, Leng Xue, et al. Fault Management in Software-Defined Networking: A Survey [J]. *IEEE Communications Surveys & Tutorials*, 2019, 21 (1): 349-392.
- [15] Lee K, Hong S, Kim S J, et al. SLAW: Self-Similar Least-Action Walk [J]. *IEEE/ACM Transactions on Networking*, 2012, 20 (2): 515–529.
- [16] Greff K, Srivastava R K, Koutník J, et al. LSTM: A Search Space Odyssey [J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2017, 28 (10): 2222-2232.
- [17] Tao Yang, Liu Xiaohu, Tao Ling. Vertical Handoff Algorithm in Heterogeneous Network Based on QoS [J]. *Computer Engineering*, 2016, 42 (12): 91-96.
- [18] Qian Zhihong, Yu Xinyi, Xu Jianhua, et al. Vertical handoff algorithm in heterogeneous networks based on vehicle terminal behavior [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2019, 49 (02): 614-623.
- [19] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks [J]. *Journal of Machine Learning Research- Proceedings Track*, 2010, 9: 249–256.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.