Computer-aided Cloud Computing for Advertising Promotion System

Xiaotian Shi[1], Yuanyuan Sun[2]

[1] Department of Art and Design, Shijiazhuang University of Applied Technology, Shijiazhuang 050035, China, shi.xt@163.com
[2] Department of Art and Design, Shijiazhuang University of Applied Technology, Shijiazhuang 050035, China, sunyuanyuansjz@126.com

Corresponding author: Xiaotian Shi, shi.xt@163.com

Abstract. With the development of technology, the profit model of Internet companies is more diversified, and more and more enterprises are gradually turning their attention from traditional media advertising to online advertising. The purpose of this research is to build a strong online advertising and process management system based on the cloud computing platform. The system enables different client-side ad delivery to be accurately and effectively controlled through a new delivery system, and automatically generates corresponding delivery plans for different types of advertisers. At the same time, the system can completely handle the authentication redirection service, cloud computing authentication service, advertisement push service and other detection and synchronization services. The system adopts a combination method of cloud computing authentication method and administrator customized push advertising URL strategy, which has practical value.

Keywords: computer aided; cloud computing; cloud advertising; intelligence; promotion
DOI: https://doi.org/10.14733/cadaps.2020.S1.90-100

1 INTRODUCTION

For existing ad serving technologies, because the delivery goals are not precise enough, the ads can't be delivered to the most likely potential audience, and users can't get the advertising information they really need. Therefore, existing ad placement is inefficient. This has caused a lot of trouble for advertisers and users: For advertisers, they spend a lot of advertising promotion costs, but they do not achieve the expected advertising effect [1]. For the user, when the user opens a website and eagerly wants to find the information he wants, he finds that the advertising information that has nothing to do with his own needs is coming, which will not only cause resentment, but also hinder users from searching and obtaining information. Therefore, the precision of advertising has become an urgent problem to be solved [2].

Yahoo's Double Click company combined user cookie analysis and banner advertising to record user online behavior through cookies, obtain user interest information, and serve advertise according to user information. This Dynamic Advertising Reporting Targeting technology is the beginning of the development of Internet advertising [3]. With the further expansion of the online advertising system, an online advertising ecosystem in which advertisements, advertisers, and other roles are used to bid on the advertising space through the advertisement exchange platform for real-time bidding and sales is formed [4].

Currently, Internet users have entered the WEB2.0 stage. Compared with the WEB1.0 era, users are more actively involved, and users are more and more involved in network activities. The wide participation of such users allows us to obtain more comprehensive and stereoscopic user data, so it is possible to analyze each different user in depth. According to statistics, Twitter's global online registered users exceeded 200 million in 2011, and online advertising revenue has exceeded 150 million US dollars [5]. Looking at the domestic market, the number of registered users of Sina Weibo has already exceeded 100 million in 2010, which is almost equivalent to half of the European population. With the popularity of broadband technology and the development and application of new technologies, the form of Internet advertising has become more diversified. At present, rich media-based Internet advertising has the advantages of being interactive, customizable, trackable, directional, creative and expressive, and evaluable [6]. The advertising marketing model has been brought into the era of digital media and has gradually formed a huge online market for computing and technology-driven. Foreign Internet giants have carried out a lot of active exploration and development in advertising technology. Microsoft has developed Content Ads, Yahoo has developed Content Match, and Google has developed Adsense. Search engine companies can use these systems to place advertisements created by advertisers on third-party website platforms [7].

Google has developed Google AdSense, an online advertising platform for medium and small businesses. Google AdSense uses computer crawlers to capture key information from different website pages to obtain key information about the website, and then use this key information to take more targeted advertising [8]. Google AdSense can place text, images, and rich media ads that match web page information on these pages based on the keywords of the page. Based on Google's powerful search engine technology and content database, Google AdSense can serve a variety of advertising formats, such as domain-based ads, search ads, and page category ads that match the content of the site, based on key information on the site. The main feature of Google AdSense is that it can serve different kinds of advertising products. Moreover, in Google's bidding ranking algorithm, due to the limitation of the ranking and level of the website, there are strict regulations on the quality of the website [9]. Due to the web-based content of Google AdSense, it is difficult for Google AdSense to obtain accurate website information when the content of the webpage is not well matched, or the labels and information are complicated. Therefore, the accurate delivery of web advertisements cannot be guaranteed. The definition of precision marketing is: as a system of communicating with customers and providing services, the system is based on precise positioning and has strong pertinence. Moreover, by relying on the system, a measurable, low-cost way of enterprise expansion can be achieved [10]. There are many means of precision marketing, which usually include Internet marketing, third-party marketing, database marketing, etc. [11]. Mejova Y et al. [12] pointed out that the effect of advertising is closely related to the relevance between web content and advertising, and for advertisements with strong competition and high relevance, the effect of advertising will be significantly improved [13]. If the content of the page is low or even irrelevant to the content of the advertisement, the effect of the advertisement will be greatly reduced [14]. With the advent of the concept of "precise marketing", the Internet advertising system began to achieve more accurate advertising by analyzing the behavior of users. According to the Boolean search condition generated by the attribute information of the user, a Boolean search is performed in the advertisement database, and a targeted advertisement conforming to the rule can be obtained. In this area, Google has developed the Google Content Network system, which uses manhole statistics to accurately target Internet ads. In addition, Google has a dedicated Demo Graphic Bidding system that is used to predict user attributes [15].

Through the user's behavioral characteristics, advertisers can deliver targeted ads to ameliorate and improve the actual performance of the ads. The research by Abdel-Basset M et al. theoretically proves this point. On the basis of collecting a large amount of user behavior data, the computer algorithm is used to judge the behavior habits of the user when accessing the product, and then the client user portrait is obtained, and the targeted advertisement is placed according to the user portrait [16]. Such an ad delivery system is based on user behavior characteristics. According to these behavioral characteristics data, matching and screening in the advertisement library to find the most relevant advertisements for delivery can not only greatly improve the user's product experience, but also have a significant effect in improving the click rate of advertisements. The study by Kim C S et al theoretically proves this point [17].

This topic proposes a cloud-based advertising push system, which can provide users with a high-quality and convenient online experience and can also set an advertisement push strategy to push advertisement pages for user terminals. Moreover, the system provides a convenient way to access the network, and push specific advertisements to users, attract users' attention, and bring more benefits to enterprises or businesses. This way of using the network to improve efficiency will certainly bring huge economic benefits to the majority of users.

2 IMPROVED FM MODEL BASED ON L-BFGS

The FM model was proposed by S. Rendle in 2010 and was used to solve the feature combination problem under sparse data. The FM model essentially introduces the quadratic parameters of the auxiliary vector to the polynomial model (this paper only considers the second-order polynomial model) to solve the problem of training polynomial model parameters in the case of data sparseness.

In the polynomial model, the combination of features x_i and x_j is generally represented by $\mathrm{x}_i\mathrm{x}_j$, that is, when $\mathrm{x}_i$and x_j are not zero, the combined feature $\mathrm{x}_i\mathrm{x}_j$ is meaningful. The second-order model equation is the formula (1):

$$\mathrm{y}(x) = w_0 + \sum_{i=1}^{F} w_i x_i + \sum_{i=1}^{F}\sum_{j=i+1}^{F} w_{ij} x_i x_j \tag{1}$$

In the formula, F represents the number of features of the sample, x_i is the value of the i-th feature, w_0, w_i, w_{ij} model parameters. The training of each quadratic parameter w_{ij} requires the use of a large number of samples with x_i and x_j being non-zero. However, in the actual recommendation system, the data sparseness is extremely high, and there are very few samples available for the polynomial model training, which leads to inaccurate parameter w_{ij} and severely affects the accuracy of the model. Based on the idea of matrix decomposition, the FM model uses a quadratic parameter w_{ij} in the polynomial model to form a symmetric matrix W. Using the matrix decomposition method, W can be decomposed into $\mathrm{W} = \mathrm{V}^T\mathrm{V}$, and the jth column of V is represented as the implicit vector of the jth dimension feature, and the quadratic parameter can be expressed as $w_{ij} = \langle v_i, v_j \rangle$. Therefore, the FM model equation is as shown in equation (2).

$$\mathrm{y}(x) = w_0 + \sum_{i=1}^{F} w_i x_i + \sum_{i=1}^{F}\sum_{j=i+1}^{F} \langle v_i, v_j \rangle x_i x_j \tag{2}$$

In the formula, v_i is the implicit vector of the i-dimensional feature, $\langle v_i, v_j \rangle$ represents the vector dot product, and the hidden vector length is $\mathrm{K}\,(\mathrm{K} << \mathrm{F})$. In formula (2), the number of parameters of the quadratic term of the FM model is $\mathrm{K} \times \mathrm{F}$, which is greatly reduced compared with the number of quadratic parameters of the polynomial model $\mathrm{n}(\mathrm{n}-1)/2$. Since w_{ij} is decomposed into dot products $\langle v_i, v_j \rangle$ of two hidden vectors, the coefficients $\langle v_k, v_i \rangle$ and $\langle v_i, v_j \rangle$of the feature combinations $x_k x_i$ and $x_i x_j$ have a common term v_i and are no longer independent of each other. That is, all samples containing a non-zero feature combination of x_i can be used to train the vector, which greatly reduces the impact of data sparsity. By simplifying the quadratic term of the FM model, the model can be rewritten as a formula (3).

$$\mathrm{y}(x) = w_0 + \sum_{i=1}^{F} w_i + \frac{1}{2}\sum_{i=1}^{K}\left(\left(\sum_{i=1}^{F} v_{i,l} x_i\right)^2 - \sum_{i=1}^{F} {v_{i,l}}^2 {x_i}^2\right) \tag{3}$$

The time complexity of the simplified FM model is optimized from O(KF2) to O(KF), which greatly improves the computational efficiency of the model.

Figure 1 illustrates an example of a feature vector for creating x under sparse data. Each row represents a feature vector of xi and a corresponding target value yi, and the middle columns represent user characteristics, movie characteristics, ratings, time, and the latest movie rating of the user, respectively. These features are combined and feature-joined, and then predicted using the FM model to obtain the user's possible score y for the movie.

Figure 1: Example of creating a feature vector of x_i.

When using the FM model for prediction, the L2 regular optimization objective function is considered, and usually the formula (4) is taken.

$$\Theta^* = arg\min_{\Theta} \sum_{i=1}^{F} \left(loss\left(\hat{y}\left(x^{(i)}\right), y^{(i)}\right) + \sum_{\theta \epsilon \Theta} \lambda_\theta \, \theta^2 \right) \tag{4}$$

In the formula, the optimization goal is to minimize the sum of sample losses under all features. Among them, $oss\left(\hat{y}\left(x^{(i)}\right), y^{(i)}\right)$ is the loss function of the i-th sample data $\left(X^{(i)}, Y^{(i)}\right)$, λ_θ represents the regularization coefficient of the parameter θ.

Rendle uses three kinds of learning algorithms, stochastic gradient descent (SGD), alternating least squares (ALS) and Markov chain Monte Carlo (MCMC), to perform parameter learning, and also has better prediction accuracy. However, in the case of a sharp increase in the amount of data, the number of iterations and computational resource consumption required by the above algorithm is high, so this paper uses a large-scale optimization algorithm more suitable for actual services: L-BFGS algorithm to perform parameter training on the FM model.

In this paper, the FM model is trained by the L-BFGS algorithm to improve the convergence speed of the algorithm. L-BFGS belongs to the quasi-Newton method: the model of the quasi-Newton method is to replace the inverse matrix H^{-1} holding the Hessian matrix with an approximate matrix B_0^{-1}. Moreover, the L-BFGS algorithm no longer stores the approximate inverse Hessian matrix B_0^{-1}, but saves the curvature information (s_k, y_k) of the T' times in the calculation process, and uses the stored curvature information to calculate the matrix H_k when needed. Since the L-BFGS stores only the curvature information of the T' number, the calculation memory consumption is greatly reduced. In the L-BFGS algorithm, the iterative formula is as shown in equation (5).

$$\mathrm{x}_{k+1} = a\mathrm{x}_k + \alpha_k P_k \tag{5}$$

In the formula, α_k is the step size, P_k is the search direction, and the formula for calculating P_k is formula (6).

$$P_k = -H_k \nabla f(x_k) \tag{6}$$

For the L-BFGS algorithm process, s_k and y_k are defined as follows.

$$s_k = x_{k+1} - x_k \tag{7}$$

$$y_k = g_{k+1} - g_k \tag{8}$$

$$\rho_k = \frac{1}{y_k^T s_k} \tag{9}$$

In the formula, $g_k = \nabla f(x_k)$. The one-dimensional array of the above formulas (7) and (8) is the curvature information of each iteration that needs to be saved.

3 DEMAND ANALYSIS AND FUNCTIONAL DESIGN

The system business includes the certification business and the advertisement push service. The following is a detailed description: The user first connects to the wireless hotspot, obtains an IP address, and then sends an HTTP packet to trigger the authentication process. After that, the AC sends messages to each other according to the OAuth protocol and the Portal server to enable users to go online. To ensure the normal operation of the system, the AC needs to periodically detect the connection with the Portal server, detect the connection with the authenticated user, and ensure the consistency with the Portal server authentication user information. Therefore, it is necessary to periodically synchronize the online user information of both parties. The administrator sends a command to the AC. When the push timing is triggered, the AC will push the advertisement URL to the authenticated user.

The functional requirements of this system are described below.

（1）Users can access addresses such as Portal server and DNS server before authentication succeeds, but they cannot access any unauthorized external network. （2）When the user authenticates, the user uses a browser to access any web address, and the AC masquerades as a destination site to establish a connection and direct the user's HTTP message to the web server. （3）The cloud portal server provides access credentials. The AC sends packets according to the credentials provided by the portal server. After receiving the complete information of the user, the user goes online. （4）After the authenticated user goes online, the access rights of the external network are obtained. The AC needs to detect whether the user is online. The user is offline according to the configuration conditions of the user. At the same time, the device maintains the link state according to the probe connection with the cloud server. （5）The AC charges the user for the flow rate, and according to the administrator configuration, it issues an advertisement URL for the user and forces the user to redirect to the advertisement page. Combined with the above description of user behavior and system requirements, it can be concluded that the system use-case analysis is as follows.

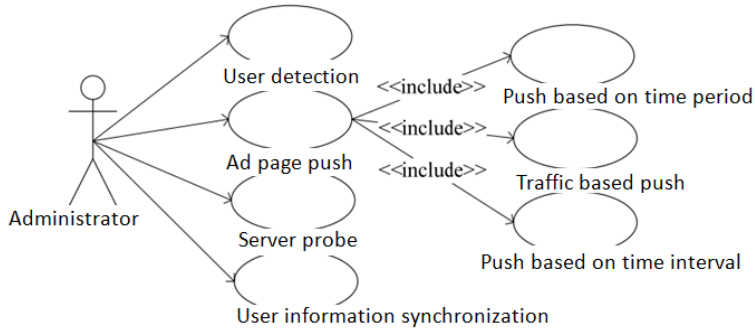

Figure 2: User use-case diagram of the push system.

The user triggers the authentication redirection function by sending an HTTP packet. The use-case of system administrator the following.

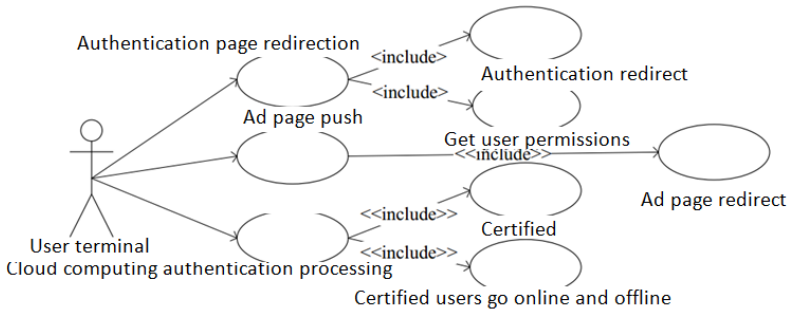

Figure 3: Push system administrator's use-case diagram.

By analyzing the functional requirements of the cloud portal authentication push system, the system needs to have the following four functions: command line processing function, cloud portal user mode function, cloud portal kernel mode function, and advertisement push function. The functions are described in detail as follows:

（1）Provide front-end interface to administrators, string processing.（2）User and server status detection, user information synchronization, and online processes of cloud Portal authentication are.（3）TCP spoofing, HTTP redirection, user authorization code, user data packet processing, and rules deliver.（4）Ad push strategy customization and ad page redirects. The functional structure of this system is shown below.

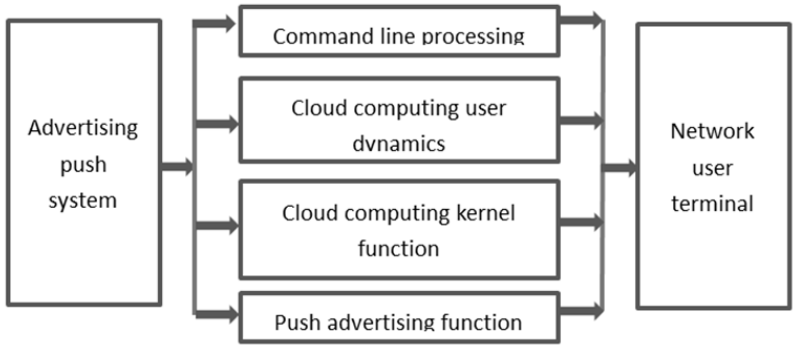

Figure 4: Structure diagram of the system function.

4 SYSTEM STRUCTURE DESIGN

The system includes Portal server detection, user detection, user synchronization, user authorization, cloud Portal authentication and authorization, and user online and offline. It is mainly divided into two parts: （1）Certification processing part: The AC sends a Code request authentication authorization. The AC sends a token credential to request the authentication user to go online. After obtaining the user information and confirming that the server receives the online confirmation packet, the AC creates a Portal user locally. The user's offline source includes: The user offline that is detected by the keep-alive mechanism and the user offline specified by the portal server （2）Related service functions: Portal server detection, user detection, and user

synchronization. These three function modules and Portal authentication services maintain the normal operation of the cloud Portal service.

This system is based on Linux and can be divided into kernel mode and user mode. The user packet is sent to the kernel mode through the forwarding chip, the kernel mode part processes the forwarding and redirection services, and the user state part is responsible for processing other fields such as the URL field in the cloud authentication service. The advertisement push part involves event triggering of user mode, rule processing of kernel mode, and modification of the rules of the switch chip. In addition, the administrator can input configuration commands to the system through the terminal. Unlike the user mode and the kernel mode, the administrator is directed to the administrator's input instead of the Internet user's message. According to the above analysis, the system is divided into the authentication kernel state, the authentication user state, the command line processing and the advertisement pushing subsystem. The system architecture can be divided as follows.

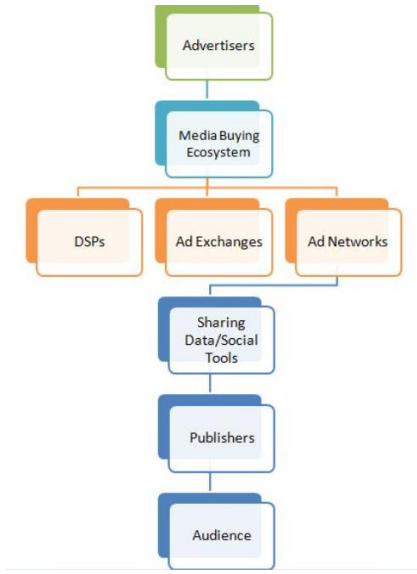

Figure 5: Structure diagram of the overall system of the system.

The external entities of the system include servers, administrators, and wireless terminals. The packet sent by the wireless terminal to the AC is first judged by the hardware switching chip. If it is not forwarded, it enters the Linux kernel state for further processing, and selects whether to send or forward directly. For the service packet of the server, the authentication processing module further processes the path name of the packet URL. The administrator's configuration commands process and deliver user interfaces through the command line interface.

The networking architecture design and service introduction of this system are described below. The system adopts a centralized WLAN networking architecture, and the AC uniformly processes the authentication service and the advertisement push service. The AC and the AP are connected through a CAPWAP tunnel, and the AP interacts with the user terminal. Combined with the above networking requirements, the network design of this system is as follows.

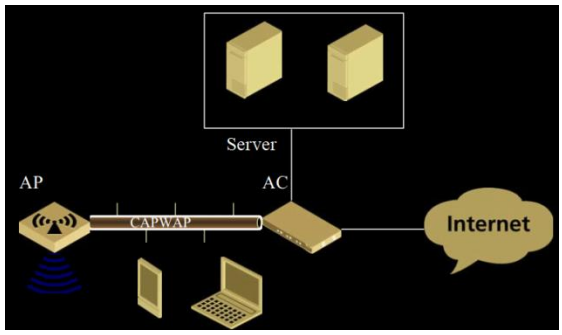

Figure 6: Networking diagram of the authentication mode.

The wireless terminal communicates with the AC layer 2, and the DHCP function is enabled on the AC to assign an IP address to the user terminal. The Cloud Portal server integrates the Web server and Portal authentication server functions. The cloud portal server sends the authentication page to the user, and sends a redirect message to the user, and then provides user information for user authentication and the AC. After the Portal authentication and authorization succeeds, the AC pushes the service-related page to the wireless terminal. The user terminal wirelessly connects to the AP, and the AP and the AC communicate through the CAPWAP tunnel. The Portal authentication service is enabled on the Vlan of the AC, and the public network IP address and external network communication are obtained through the interface. The user only needs to use the IP address assigned for the first time.

The OAuth protocol specifies the following entities: the client (which refers to the AC in this system), the authorization server (which is the Portal server in this system), service provider (which is referred to as a portal server in this system) and a resource owner (which in this system refers to a wireless access terminal, that is, a user). The agreement stipulates that the service provider provides the client with an authorization layer. The client does not have the authority to use the resources of the service provider's server and can only use the authorization code to invoke the interface of the authorization layer. After the client logs in to the authorization layer, it can access some resources that the resource owner opens to the terminal according to the permissions of the authorization layer. The protocol implements the function of user-specified open resources by means of open partial authorization. The protocol running sequence is as follows.

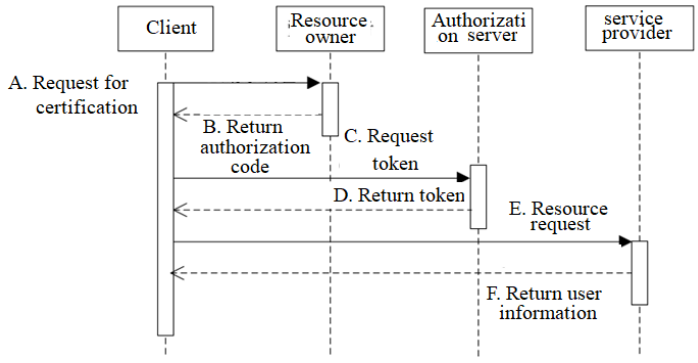

Figure 7: Sequence diagram of the OAuth protocol operation.

After the cloud portal authentication is successful, the AC needs to create a Portal user locally. The method of cloud Portal authentication is PAP authentication. The PAP authentication mode indicates that the username and password used in the authentication process are plain text. Moreover, in the network design of the system, the user is connected to the second layer of the AC. Therefore, the AC can directly obtain the MAC address and IP address of the user. After the user is successfully authenticated, the user rules delivered by the AC include the IP address and the MAC address. After the AC obtains the user information through the cloud Portal server, the portal user completes the online user online process. The sequence diagram is as follows:

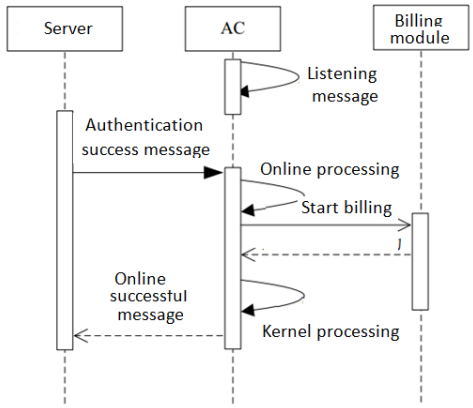

Figure 8: Sequence diagram of the certification online.

(1) The AC device opens the port and listens to the HTTP packets of the Portal server, parses the packet URL, and matches the online service field. After the user receives the cloud authentication success message, the AC adds the authentication user to the authentication user chain to process subsequent packet forwarding and rule delivery. (2) The device creates a Portal authentication user and extracts the authentication and authorization information sent by the cloud portal server. The forwarding processing part is notified by the device to add a user rule in the kernel mode, and the source IP and source MAC of the rule are the user's IP and MAC. At the same time, the user rule is sent to the driver, and the user data packet is subsequently forwarded by the hardware driver. Then, traffic is started to be counted for the authenticated user.

5 CONCLUSION

This thesis first studied the development of cloud computing authentication push technology, analyzed related technologies, and proposes a cloud-based authentication push scheme to design this system, which is innovative. The system adopts the combination of cloud computing authentication mode and administrator customized push advertising URL strategy, which has practical value. Through the design and implementation of the system, we can draw the following conclusions: (1) The system is stable and related functions are well implemented. The method of session hijacking is used to redirect users and further perform authentication services. This method is widely used in the industry and is equally applicable in this system. (2) The advertisement URL is pushed to the user by modifying the driver and kernel related rules, which is suitable for the service design of the switch. (3) The system implements the authentication and push service functions. Compared with the implementation of the traditional authentication system, the innovation of this system is to realize the authentication method based on cloud computing address. (4) The system adopts the OAuth protocol interaction, which is different from the traditional Portal

authentication protocol. It supports the third-party account authentication mode naturally, and the subsequent expansion is more convenient.

6 ORCID

Xiaotian Shi, https://orcid.org/0000-0003-0233-3530
Yuanyuan Sun, https://orcid.org/0000-0002-4577-3854

REFERENCES

[1] Yang, K. C.; et al: Consumer attitudes toward online video advertisement: YouTube as a platform, Kybernetes, 46(5), 2017, K-03-2016-0038, https://doi.org/10.1108/K-03-2016-0038
[2] Gauba, H.; et al: Prediction of advertisement preference by fusing EEG response and sentiment analysis, Neural Networks, 2017, S0893608017300345, https://doi.org/10.1016/j.neunet.2017.01.013
[3] Reinhold, B.; et al: Associations of attitudes towards electronic cigarettes with advertisement exposure and social determinants: a cross sectional study, Tobacco Induced Diseases, 15(1), 2017, https://doi.org/10.1186/s12971-017-0118-y
[4] Ray, A.; et al: Creative tagline generation framework for product advertisement, IBM Journal of Research and Development, 2019, 1-1, https://doi.org/10.1147/JRD.2019.2893900
[5] Tully, M.; Vraga, E. K.: Effectiveness of a news media literacy advertisement in partisan versus nonpartisan online media contexts, Journal of Broadcasting & Electronic Media, 61(1), 2017, 144-162, https://doi.org/10.1080/08838151.2016.1273923
[6] Thi, N.; Che-Pin, C.; Shyan-Ming, Y.: A Wi-fi union mechanism for internet advertising reciprocal platform in microenterprises, Sensors, 17(7), 2017, 1617-, https://doi.org/10.3390/s17071617
[7] Kille, J.: A content analysis of health and safety communications among internet-based sex work advertisements: important information for public health, Journal of Medical Internet Research, 19(4), 2017, e111, https://doi.org/10.2196/jmir.6746
[8] Saha, K.; et al: Characterizing awareness of schizophrenia among Facebook users by leveraging Facebook advertisement estimates, Journal of Medical Internet Research, 19(5), 2017, e156, https://doi.org/10.2196/jmir.6815
[9] Einziger, G.; Chiasserini, C. F.; Malandrino, F.: Scheduling advertisement delivery in vehicular networks, IEEE Transactions on Mobile Computing, PP(99), 2018, 1-1, https://doi.org/10.1109/TMC.2018.2829517
[10] Yin, L.; et al: A game-theoretic approach to advertisement dissemination in ephemeral networks, World Wide Web-internet & Web Information Systems, 2018, https://doi.org/10.1007/s11280-017-0432-6
[11] Bajaj, S.: Food safety in the 21st century: Regulation of advertisement for food products in India — advertisement for food products, Food Safety in Century, 2017, 469-477, https://doi.org/10.1016/B978-0-12-801773-9.00038-8
[12] Mejova, Y.; Weber, I.; Fernandezluque, L.: Online health monitoring using Facebook advertisement audience estimates in the United States: Evaluation study, Jmir Public Health & Surveillance, 4(1), 2018, e30, https://doi.org/10.2196/publichealth.7217
[13] Choi, I.; et al: Using different Facebook advertisements to recruit men for an online mental health study: Engagement and selection bias, Internet Interventions, 8, 2017, 27-34, https://doi.org/10.1016/j.invent.2017.02.002
[14] None. IEEE Internet Computing: Call for Papers House Advertisement, IEEE Internet Computing, 21(2), 2017, c3-c3, https://doi.org/10.1109/mic.2017.32
[15] O'Meara, G.: Mining and classifying images from an advertisement image remover, Annals of Data Science, 2018, https://doi.org/10.1007/s40745-018-0164-1

[16] Aikin, K. J.; et al: Correction of misleading information in prescription drug television advertising: The roles of advertisement similarity and time delay, Research in Social & Administrative Pharmacy Rsap, 13(2), 2017, 378-388, https://doi.org/10.1016/j.sapharm.2016.04.004

[17] Kim, C. S.; Lee, M. S.; Park, C. H.: Implementation of an app scheduler for the effective display of advertisement contents on android platform, Psychiatrische Praxis, 12(11), 2018, 129–134, https://doi.org/10.1055/s-0044-101530