Design And Implement Of TCP/IP Protocol Monitor Submodule In Three Screen Protocol Adaption System

Hong Lin
Network and Information Center
North China Electric Power University
Beijing, China
e-mail: linh@ncepu.edu.cn

Yajuan Sun
Network and Information Center
North China Electric Power University
Beijing, China
e-mail: syj@ncepu.edu.cn

Baohui Wang
Software School
Beihang University
Beijing, China
e-mail: wangbh@buaa.edu.cn

Zhirou Zhang
Network and Information Center
North China Electric Power University
Beijing, China
e-mail:zhangzr@ncepu.edu.cn

***Abstract*—Because of the integration of three networks in China, search requirements with multimedia information will increase. The contents on TV, computer and cell phone are different in format and protocol not to intercommunication and share. When the national policy demand makes the division up, the content provider must provide formats and adapter mechanisms which are intercommunicated. As precondition of Three Screen Adaption, TCP/IP PROTOCOL Monitoring's accuracy determines the effect of three screen fusion directly. In this paper, the methods of protocol monitoring are introduced firstly. With comparing, the existing monitoring's accuracy can't meet the requirement of three screen protocol adaption. Then a new protocol monitoring model in three screen protocol adaption environment is proposed and protocol recognizer is designed and optimized with HMM technique. It will simplify development and publishing information unified that three screen equipment access the same protocol monitoring model, which has important practical value and application prospect.**

Keywords-Protocol Monitoring; Protocol Recognition; HMM; Three Screen Fusion

I. INTRODUCTION

With the development of search platform on the Internet, search engine has become important distributed platform of network media and information, instead of basic tool. And allow for mobile networks and Next Generation Networks (NGN) that are based on three network fusion, search engine should support three screen fusion services of cell phone, computer, Television. "Three screen fusion" means making full use of existing search engine platform and resources to form the complementary transfer of video information and unified services between TV screen, cell phone screen and computer screen, all of which are user-centered, to increase the value of search engine in three screen area. So it is critical to solve the differences between 3 kind of terminal service interaction protocols of phone, and computer television, and shield the heterogeneity of these terminals. Research emphasis of protocol adaption is making different terminals get the same-effect information.

The author National joined in sub-project "Three Screen Fusion Service Platform Supporting Enhanced Search Engine" (Project No. 2011BAH11B04) of National Science and Technology Support Program "Key Technology Research and Demonstration for the Enhanced Search Engine". In this sub-project, a three screen fusion service platform supporting enhanced search engine need to be established. TCP/IP PROTOCOL monitoring is precondition for protocol transmission and protocol data transfer. Only when the type of protocol in the specific network has been recognized and the protocol has been analyzed, characters of the network including information structure, network topology, routing algorithm, IP distribution can be obtained, and protocol can be adapted to implement three network fusion. Traditional protocol monitor mechanisms have too little accuracy to meet the requirement of three screen protocol adaption. So accuracy of TCP/IP PROTOCOL monitoring is emphasis and difficulty in the research of this project.

II. REQUIREMENT ANALYSIS OF PROTOCOL MONITORING IN THREE SCREEN PROTOCOL ADAPTION

Enhanced Search Engine, of which Protocol monitoring is a sub-function, is used in open, melting, complex network environment that come from three network fusion. So the requirement to protocol monitoring is monitoring as accurate as possible with the least cost.

The targets of protocol adaption project are to solve communication protocol difference between main three screen equipment, and implement protocol monitor, analysis and intelligent protocol adaption to meet the business requirement of three screen users. When users can access services with terminals (cell phone, computer, or digital TV), the user requests are monitored, analyzed and adapted to protocols according to different types of requests, by services. Adapted request data is send to a specific protocol to enquiry and the enquiry function transmits the request to relative portal system to finish seamless connection of requests. Access data monitoring are divided into 3 parts: HTTP protocol data monitoring, WAP protocol data monitoring, HTTP protocol control command data monitoring. [1, 2]

Business flow of protocol monitoring service has several branches. Different threads need to be established to process data of 3 kinds of protocols respectively. HTTP protocol data monitoring and WAP protocol data monitoring process request data from terminals, and HTTP protocol control command data monitoring process multimedia control commands from search engine media control, so the size of request data packet and multimedia control command data packet are not large, but there are a large number of concurrent request and the time of concurrency is short [3].

III. Protocol Monitor Method and Model Design

A. *Comparison of Protocol Monitor Methods*

Some protocol monitor techniques are analyzed; advantages and disadvantages of main protocol monitor methods are showed in Tab 1.

TABLE I. Comparison of Main Protocol Monitor Method

Protocol Monitor Method	**Accuracy**	**Real-time**	**Performance**	**Expansibility**	**Division**	**encryption**
Port Protocol Monitor	Low	Good	Less high	weaker	Bad	Weaker
Characteristic String Match Protocol Monitor	High	Fine	Low	weak	Good	Weak
Traffic Fearture Protocol Monitor	lower	Bad	High	Strong	Bad	Strong

With the analysis in Tab. 1, port protocol monitoring, as basic, simple, and high-efficient method, can't meet the requirement of networks now. Because feature library become larger and larger and time-space complexity of system become larger and larger, the efficiency of the monitor technique based on characteristic string become lower and lower, the monitor technique based on traffic feature can't recognize the kind of encryption protocol accurately. Characteristic string monitoring technique can implement accurate characteristic match and misjudge rarely, adapt to complex network environment to process problems of packet loss and recombination. This technique is most general monitor method now. The monitor technique based on traffic feature is bad real-time, so it is less used in practice.

B. *Protocol Monitor Model Design*

In three screen adaption system, when request data is monitored, monitor recognize corresponding protocol for protocol adaption, and then push it to protocol adaption module to finish the three screen protocol adaption workflow. So a protocol monitoring model meeting the requirement of three screen protocol adaption should be established, which can not only recognize traditional specific port protocol but also recognize new unspecific port protocol that is difficult to recognize. The protocol monitor system is design based on this model.

There are the following rules for this protocol monitoring model:

- Simple and Practical

Complex model is difficult to use, on the contrary, simple model is easy to understand and use.

- Sufficiency is Enough

Protocol monitor model is served for protocol adaption design, insufficiency and excess in system design should be avoided. So insufficiency and excess in model design also be avoided.

- Comprehensively Considering Accuracy, Efficiency and Cost

Different kind of protocols needs different protocol monitor techniques. To recognize all kinds of access protocols, multiple protocol monitor techniques are needed comprehensively, and the cost of system also be considered.

- Adaptability

Protocol monitor model should adapt to some change factors. Presupposition and environment's change have large effect to monitor system, which means this system is fragile and the model is impractical.

- Generality

As a model, this model should be general, not to some specific system or productions.

Current monitor model can't meet the requirement of three screen protocol adaption system which need comprehensively use multiple methods and techniques. A monitor model for three screen protocol adaption is proposed, see also Fig. 1.

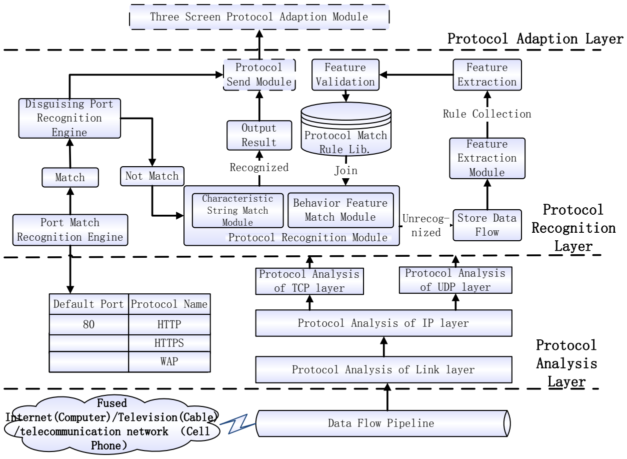
Three Screen Protocol Adaption Module
Protocol Adaption Layer
Protocol Send Module
Feature Validation
Feature Extraction
Disguising Port Recognition Engine
Output Result
Protocol Match Rule Lib.
Rule Collection
Feature Extraction Module
Recognized
Join
Match
Not Match
Characteristic String Match Module
Behavior Feature Match Module
Protocol Recognition Module
Unrecog-nized
Store Data Flow
Port Match Recognition Engine
Protocol Recognition Layer
Protocol Analysis of TCP layer
Protocol Analysis of UDP layer
Default Port
Protocol Name
80
HTTP
HTTPS
WAP
Protocol Analysis of IP layer
Protocol Analysis of Link layer
Protocol Analysis Layer
Fused Internet(Computer)/Television(Cable)/telecommunication network (Cell Phone)
Data Flow Pipeline

Figure 1. Protocol Monitor Model of Three screen protocol Adaption

Now this protocol monitor model is introduced with functions. The model are composed of 3 layers from top to bottom, which are Protocol Adaption layer, Protocol Recognition layer, Protocol Analysis layer respectively.

Protocol Analysis layer. Users access search engine with computer, cellphone, TV. TCP/IP PROTOCOL is analyzed, then the packet header and key payload information are provided to Protocol Recognition layer in the form of flow [4].

Protocol Recognition layer. It is a kernel of the model, which provides all recognition functions with expansible protocol recognition engine. With dual-core Protocol Recognition module, dynamic behavior feature match of comprehensive flow, characteristic string match of load and default port load are implemented to recognize protocol and get protocol feature of each flow. These obtained protocol information, protocol data, protocol recognition information are sent to protocol adaption module. Port match recognition engine recognize general protocol which use specific port. Source port and destination port are obtained from the flow, then ports are recognized with general port mapping table and the disguising ports are abandoned.

Protocol Adaption Layer. Its main functions are protocol adaption, syntactic structure extraction, reading cache files and judge request data's type. If the data is HTTP protocol data, then it returns relative data to terminal directly. If the data is WAP data, then it transmits WAP to HTTP. If the data is SIP protocol control command, then it transmits SIP to HTTP [5].

C. *Protocol Monitor Recognizer Design*

Presently, port protocol monitor technique and characteristic string monitor technique are used for protocol monitor mostly, whose precondition is RFC is obeyed strictly. But with appearance of new protocols that are not using a specific port but using ports with dynamic negotiation in the course of running, these traditional techniques act more and more limitedly. In three screen protocol adaption system, for fast and high-efficient protocol monitoring, protocol Feature Library which have higher match rules should be establish with less data, and encryption protocol should be recognized to accuracy to protocol monitoring.

Protocol recognition technique of Hidden Markov Model(HMM) can meet the above requirements, which can establish an almost complete Protocol Feature Lib with less training data and the feature library is much smaller than that with other techniques so as to recognize protocols quickly.

The design and implement of protocol monitor module mostly meet function requirements of above analysis, but its accuracy is too low to meet the relative requirements of three screen protocol adaption. The author proposed a idea of recognizer based on dual features, according to the analysis of recognition accuracy of traditional single feature recognizer, and compared their monitoring accuracies.

Given a specific observation sequence O, set C (K kind altogether), and model $\lambda = \lambda 1, \lambda 2, ..., \lambda K$, recognizer computes which kind have largest probability to this observation sequence in $c \in C$, i.e. c=class(O). In this paper, 2 kind of recognizers are analyzed. One is recognizing with Maximum Likelihood algorithm, i.e. class(O)=argmaxc (O|λc), in which argmaxc is class c that get Maximum Likelihood value in C. The other uses Viterbi algorithm to find most possible state sequence, whose probability is P(O,λ)maxP(O,S|λ)viterbiS. Given a output sequence O, Viterbi classifier get Viterbi route of every model λi 's sequence and choose the model that have optimal route, whose math expression is class(O)=argmaxc Pviterbi(O,λc).

Comparing with WAP, HTTP have higher recognition ratio, and SIP have lower recognition ratio [6]. SIP has behavior mode and its features can't be generalized completely, so its recognition effect is not good.

With vector quantification technique, recognition with 2 features together can be implemented. Vector quantification technique is that given the data <arrival time, size> of every packet, time is logarithmically transformed to reduce dynamic range, time and size are given the same weight value, <logarithm of arrival time, size of packet> is normalized to <-1, +1>. Thus, bidimensional packet data are transformed to unidimensional to have the information of size and time processed at the same time to improve recognition performance of model.

HMM model is divided into 2 kind with its transfer directions: from client to server and from server to client. K-means clustering algorithm is used for every vector independently to recognize a set of data collection packets of representative carrier or code word. N code words are quantized with code book to a random cluster centered on k = N / 2 vector, and according to vector <arrival time, size>, the most proximate clustering distribution vector is obtained in every iteration. With the result, size of packet and time information is combined into a model and protocol recognition ratio will be improved.

IV. ARCHITECTRUE DESIGN OF PROTOCOL MONITORING SYSTEM

In logical architecture, protocol monitor are divided into 5 parts: three screen clients, protocol monitoring layer, protocol adaption layer, log process layer, configuration and control layer [7], see also Fig 2.

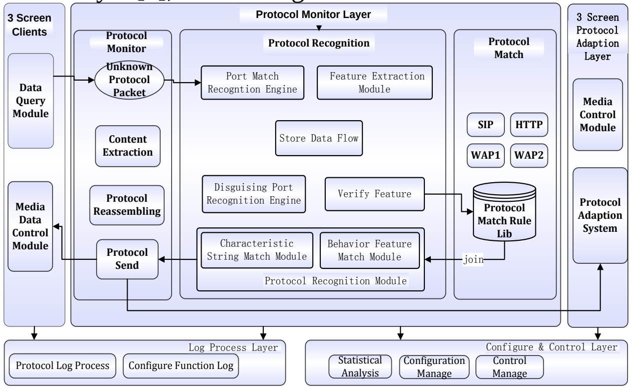

Figure 2. System Logical Architecture Diagram

When three screen client access the service system, the system send it to corresponding protocol monitor layer according to the protocol type of this request with user protocol reassembling and analysis, then monitored protocol is matched for Protocol Match Rule Library to recognize. Recognized protocol is sent to protocol adaption layer, and transformed and forwarded to corresponding portal system by this layer to finish the seamless connection of this request. When users access flow media resource with terminal equipment, protocol monitor layer, protocol monitor layer send the request to service provision system according to the type of terminal request protocol. Service provision system connect user for session with specific transmission protocol. Log process layer and configuration and control layer control and manage protocol monitor layer and protocol adaption layer.

V. Conclusions

Accuracy ratio analyses of the protocol monitor models optimized and not-optimized show that the protocol monitor model of three screen protocol adaption system proposed in this paper has higher monitor success ratio and accuracy ratio. Because of limited research time, only HTTP, SIP, WAP, HTTPS are tested. For some encryption protocol monitor, recognition success ratio is also higher when protocol features are described correctly.

Researching TCP/IP protocol monitor, quickly recognizing and monitoring all kinds of protocols, keeping integrity of session course between different protocols, making SIP equipment read and operate protocol monitor model that is support HTTP, are to have three screen equipment access the same protocol monitor model, simplify development and release information unified, which has important practical value and popular prospect.

Reference

[1] D. Jacobson, Introdution to Network Security, Beijing: China Electronic Industry Press, 2011:123-127

[2] Behrouz A.Forouzan，TCP/IP Protocol Suite (4th Edition), Beijing: China Tsinghua University Press, 2011：47-65

[3] Songlin Kang, Sihang Li, Intrusion Detection System Design and Implementation Based on Network, Chinese Journal of Computer and Information Technology, No.19, 2011:1-3

[4] Wenlong Xiao, Songru Lin, TCP/IP Best Introduction (6th Edition), Beijing: China Machine Press , 2010

[5] Charles Kozierok, The TCP/IP-Guide: A Comprehensive, Illustrated Internet Protocols Reference (Volumn 1). Beijing: China Post &Telecom Press. 2008

[6] D.E. Comer, Internetworking with TCP/IP: Principles, protocols, and architecture (5th Edition). Beijing: China Electronic Industry Press. 2007

[7] Weiwei Wang, Xuefeng Zheng, Technology of sniffer and anti-sniffer in LAN, Chinese Journal of Computer Engineering and Design, 2005,26(11) :30-56