Software Requirement Analysis

7 November 2005

SmartGuardian

SmartGuardian™ Software

SmartSoft Network Solutions, Inc.

Tayfun Şen - Özden Meren - Sedat Zelyüt - Hakan Özadam - Akın Öztürk

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Purpose

This document describes the software requirements of the SmartGuardian project. It aims to help the developers to clearly understand what the customers wish to obtain from the product. The complete description of the functions and the external interfaces of the product are specified in this report. It includes the performance specifications, and information about development.

This document is prepared to serve software developers and designers as a guideline in developing and testing the software. It is designed to help the users and customers understand what SmartGuardian is supposed to achive.

1.2 Scope

In the information technology age today, one of the crucial issues is security. Every computer needs to be secured but also the whole organization network should be protected against attackers and outside threats. Moreover corporations have limited network resources for Internet access, therefore the usage should be handled efficiently and fairly. Furthermore, to improve organization efficiency, internet access should be restricted according to some policies. This will ensure that the employees only use the Internet for work related purposes. Our project SmartGuardian aims to answer all the issues mentioned above.

1.2.1 Software Products

The software products we are planning to produce are: SmartGuardian Wall, SmartGuardian Crawler, SmartGuardian Controller and SmartGuardian Reporter.

SmartGuardian Wall will be responsible for security and filtering issues. First of all, it will be a firewall protecting the system from malicious codes, trojans, worms, and viruses. It will control programs' access to the external world and also it will protect the system from outside threats. This protection will be provided in the application level by inspecting packet contents of both outgoing and incoming data. In addition, the program will be able to forbid protocols such as ftp, smtp, http etc. Moreover the program will block ports and IPs determined by the network administrator. Secondly SmartGuardian Wall will handle filtering issues. It will apply content filtering to the web sites requested by the user. According to the rules and policies defined by the administrators, users will be able to access these web sites or not.

SmartGuardian Crawler will use data mining and AI methods classify new web sites that have no prior records in the database. It will look for keywords and apply data mining techniques in the html files and use these to classify web sites so that the SmartGuardian Wall can use these classes. The websites are categorized into groups such as Pornography/Nudity, Entertainment, Shopping, Games/Gambling, etc.

SmartGuardian Reporter is a tool to help the administrator to analyze the system. All actions taken by the SmartGuardian Wall and all attacks to the system will be logged, and the reporter will display these data in a textual and graphical manner. The administrator will also be able to change view settings to see different statistics and/or in a different appearance.

SmartGuardian Controller is an interface that will be used to control SmartGuardian Wall and SmartGuardian Crawler. Using the Controller the administrator will be able to assign which programs will have rights to access the internet, which protocols are allowed and so on. Also the administrator will be able to control domain classes and he will be able to add/remove domains from these lists.

1.2.2 Software Applications

Our software is planned to be used in all kinds of companies and government organizations such as: schools, libraries, administrative institutions and commercial companies of all sizes. The main benefits of using our software is increasing employee efficiency and removing internet threats. According to a recent research [1], half of the time the employees spend on the Internet is unrelated with work and can be considered waste of time. And in general, the companies spend a huge resource in terms of money and time in maintaining their network systems security. The main threats stem from careless employees who use instant messaging, P2P and who access dangerous sites which contain malicious code. SmartGuardian will also enable libraries and schools to focus on

educational and research objectives. We are working for you to make Internet a safer place for all.

1.3 Market Research

We have done market research to better understand the requirements to be addressed by our program. We also found out about what the most used programs are and their shortcomings in terms of usage and functionality. We visited some companies in the Technocity area in METU and also the computer center in the university campus. The whole list of organizations we have visited is presented next.

1. Computer Center
2. Milsoft
3. TR.NET
4. Aselsan
5. Aydın Yazılım
6. Halıcı Yazılımevi

Some companies were very helpful in answering our survey, others were reckless. Here are the questions used in our survey.

Are you using a gateway currently? If yes,

1. Which database software are you using?
2. On which OS is it running?
3. Which web server does it work with?
4. How stable is your gateway?
5. Which problems have you faced up to this day?

6. Are there any features which does not exist in your gateway and you wished to have?

Which statistical information would you like to be kept by the gateway?

7. URL hits?
8. Number of times a server has been accessed and when it has been accessed?
9. Bandwidth usage of the users?
10. Would you like to have different groups to have different permissions?

What kind of hardware would you be willing to spare for this product?

11. CPU
12. RAM
13. HDD

What are your performance expectations for a gateway?

14. To how many concurrent users should the gateway be able to handle? In other words, how many people will use this gateway in your organization?
15. Do you have any other comments that you would like to add?

Here are the results of our survey:

SURVEY RESULTS

Question No	Aydın Yazılım	Aselsan	Milsoft	Halıcı Yazılım
1	XML	MS SQL Server	Text file	SQL
2	Windows Server 2003	Windows Server 2003	Slackware	Windows Server 2003

3	Apache	-	Apache	IIS
4	-	-	-	Hardware problems
5	-	-	-	-
6	Advanced Reporting	Efficiency	Real-time monitoring	Antivirus functionality
7	Yes	Yes	Yes	Yes
8	Yes	Yes	Yes	Yes
9	Yes	Yes	Yes	Yes
10	Domain-wise	Complete	Complete	Complete
11	P4	P4 2 GHz Xeon	P3 1 GHz	P4
12	512 MB	-	1 GB	1 GB
13	80 GB	-	60 GB	30 GB
14	200	1500	700	40

1.4 Definitions, Acronyms and Abbreviations

1.4.1 Definitions

Domain Classes: A domain class is a set of domains that are logically categorized according to their content similarity.

Policy: A policy is a set of rules that works together with a shared objective.

Rule: A rule is a principle deciding whether a given request has permission to access the Internet.

1.4.2 Acronyms & Abbreviations

AI: Artificial Intelligence

DFD: Data Flow Diagram

IEEE: Institute of Electrical and Electronics Engineers

LAN: Local Area Network

SRS: Software Requirement Specifications

Std.: Standard

1.5 References

[1] See secure computing white paper on:

http://www.securecomputing.co.kr/scc_bestpractices.pdf

2. IEEE Recommended Practice for Software Requirements Specifications (IEEE std 830-1998). This standard can be obtained from the IEEE web site at:

http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=15571

2 OVERALL DESCRIPTION

2.1 Product Perspective

2.1.1 System Interface

The system interface of our software product can be seen from the DFD diagram in figure 1. As the figure implies, SmartGuardian has three types of interfaces. The first type of interface is with the LAN where the client computers connect to SmartGuardian server for permission to connect to the Internet. Requests are handled by SmartGuardian and according to the policies a verified response is transferred back to the client. The second interface is the utilized when SmartGuardian connects with the Internet. Client generated and verified requests or requests generated by the SmartGuardian itself are transferred to the Internet. A final interface is utilized by the administrator for setting up SmartGuardian and for SmartGuardian to return feedback.

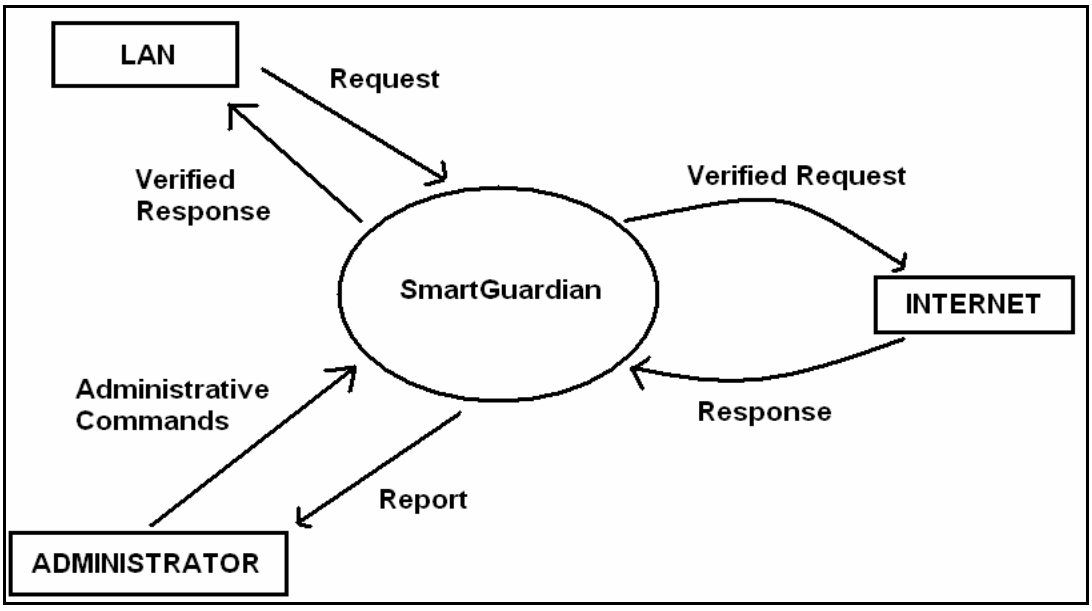

Figure 1: Context – Level DFD (Level 0)

2.1.2 User Interfaces

There will be a web based user interface through which the network administrator has complete control over the system. The administrator will be able to create new rules, policies and user groups. Access to user interface requires authentication with an admistrator login. The interface shall display current status of the whole system. This user interface is called SmartGuardian Controller which has four menus in order to administrate the modules systematically.

First menu is about general system configuration. These configurations are TCP/IP settings of the network which the gateway is serving, switches which activate or inactivate the modules of the system.

Second menu is for user setup which involves adding and deleting users and defining user groups, all other information about users can be viewed and modified through this interface.

Third menu is for configuring SmartGuardian Wall. State of all ports can be set by this menu. Also IP blocking and restriction settings can be done in this part of the menu. All other changes to firewall policy and bandwidth limits can also be performed though this menu.

Fourth menu is for configuring traffic filtering. The rules of content filtering, file downloads, URL & Domain access and the allowed applications which use internet sources are defined here. For content filtering URL and domains are classified into groups. The administrator can add and delete those groups or add

and delete URLs and domains to those groups. Administrator can define keywords and extensions which are used to restrict file downloads.

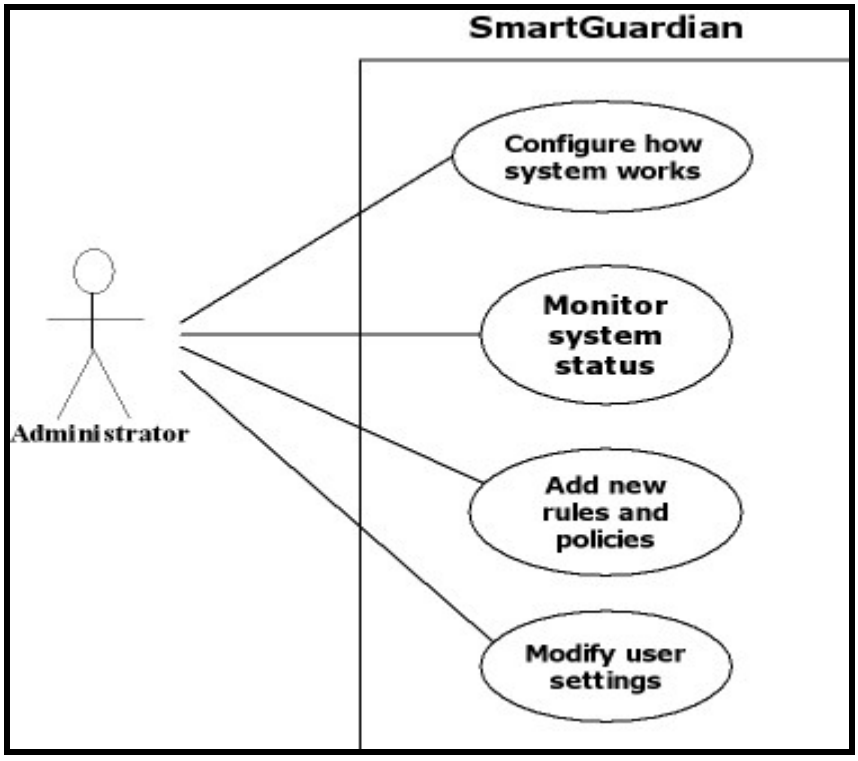

figure 2: Use case diagram for the administrator

The interface which ordinary users will use is going to be transparent to them. Thus, the ordinary users not going to be bothered by any technical management whatsoever. They will connect to Internet through SmartGuardian software and this will be handled automatically by the system.

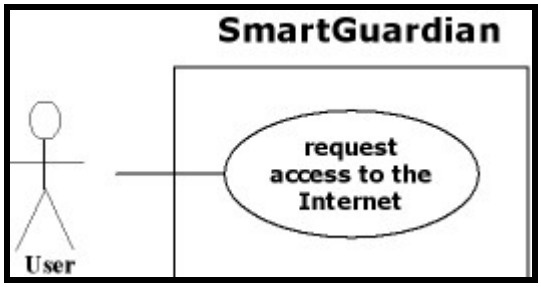

figure 3: Use case diagram for the normal user

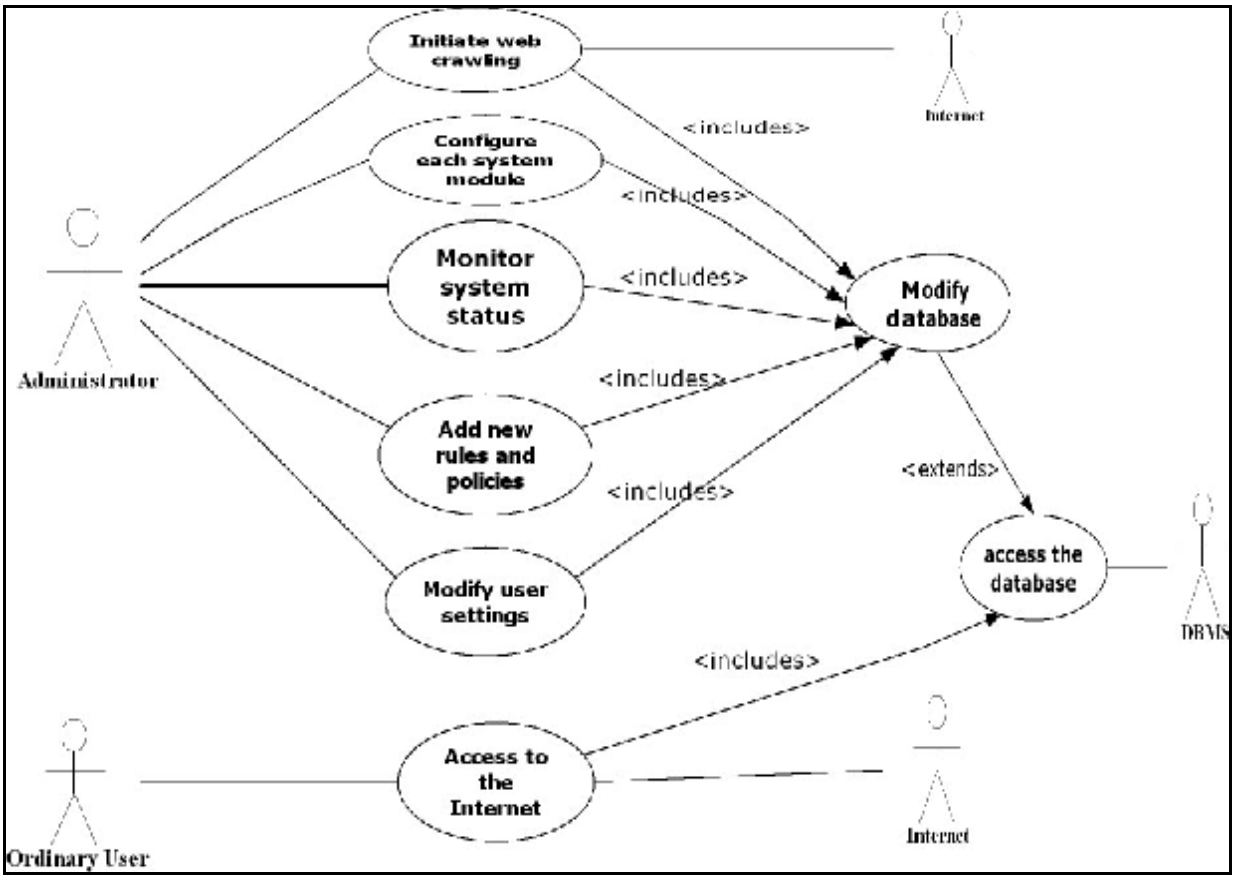

figure 4: The general use case diagram

2.1.3 Software Interfaces

The operating system that SmartGuardian works on is FEDORA CORE 4. SmartGuardian requires MySQL 4.1 for database applications. The user interface will be implemented in PHP 4.0 which runs on APACHE 2.1 Web Server. Both SmartGuardian Wall and Reporter will use database directly in order to get configuration settings. PHP will also communicate with MySQL database in order to display user interface.

2.1.4 Memory Constraints

Memory constraints depend on the size of the network the system is serving. For maximum functionality and a network consisting of 1000 users we specify 1 Gigabytes of main memory and 40 Gigabytes secondary memory.

2.1.5 Operations

The system depends on a data mining process which is both IO and CPU bounded operation. Therefore, this data mining process should run on time intervals set by the network administrator when network congestion is low. There is also one subroutine which is supposed to run regularly to keep system backups.

2.2 Product Functions

The major tasks SmartGuardian is expected to achieve are web filtering, security and monitoring.

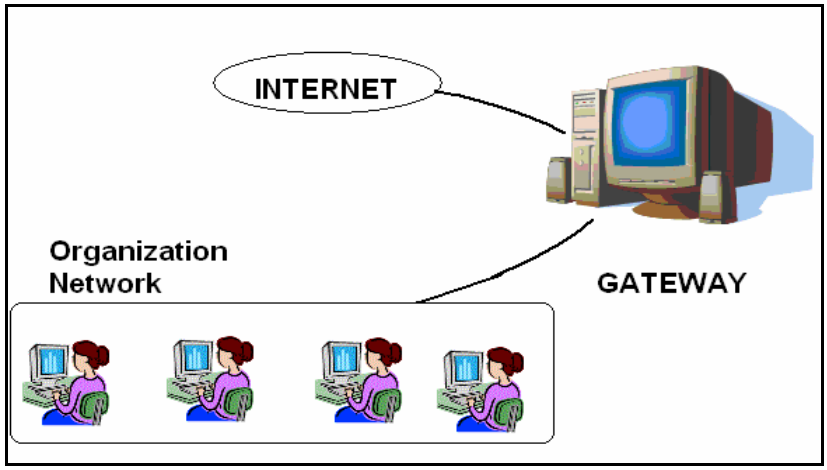

figure 5: SmartGuardian's place in the network.

SmartGuardian Wall will provide security, reliability and balanced distribution of bandwidth. It will manage Internet access of all programs that are running on client machines in the view of the rules that are set by the network administrator. Another function of the Wall is to provide user groups balanced bandwidth consumption. Still another function of the SmartGuardian Wall is web filtering. It will filter the incoming and outgoing network transactions according to their content, URL, domain and IP. The administrator shall be able to restrict or allow specific domains and URLs that can be accessed by the users of the local area network. It will restrict file downloads according to their extensions and predetermined keyword in the name of the files. It will also restrict the internet access of users in terms of time and other classifications.

SmartGuardian Crawler will handle data mining which is crucial for the filtering purposes. It will periodically search the web pages that have been previously accessed and will categorize them according to their content.

SmartGuardian Reporter handles the monitoring and logging tasks in various ways for administrative purposes. It will monitor the network traffic; show the statistical data of the usage of network resources; keep detailed logs of users.

SmartGuardian Controller is the interface of the system; the administrator can make changes to the settings of the other modules.

2.3 User Characteristics

The expected user profile of SmartGuardian is network administrators or those who have some background in computer communication and networks.

2.4 Constraints

Reliability & Criticality; Internet access is crucial for all companies so any failure in Internet access results in considerable time and money loss. That is why stability is an indispensable characteristic of SmartGuardian.

Safety & Security; Internet is an important threat to all computers which are connected to it therefore one of the most important objectives of SmartGuardian is to bring security to its local area network.

2.5 Development Process

It is clear that some functionalities of SmartGuardian are indispensable while some of them are optional and we have a tight schedule. Therefore we want to be able to test and modify the code in every step of project development. That is why we will develop SmartGuardian in an evolutionary aspect. We will adopt the incremental process model and apply it in our project development.

3 SPECIFIC REQUIREMENTS

The product will be composed of four main parts, namely SmartGuardian Wall, SmartGuardian Crawler, SmartGuardian Reporter and SmartGuardian Controller. Below are their detailed descriptions.

3.1 SmartGuardian Wall

SmartGuardian Wall will have a firewall functionality as a primary protection against all attacks and malicious content. Since our gateway is a point of access to the Internet for a lot of computers in a company, it is very crucial to offer a high level of protection to these client computers. SmartGuardian Wall will also be responsible for web filtering and control of the web traffic. It will limit the internet access of the clients according to the rules defined by administrator for determining valid content. SmartGuardian Wall uses policies consisting of rules to operate, so to carry on our description, we are now presenting rules.

3.1.1 Rules

There are eight different types of rules concerning the Wall. These are:

a) Application Rules

b) Protocol Rules

c) Port Rules

d) Bandwidth Rules

e) Download Rules

f) File Download Operation Rules

g) URL/Domain/IP Access Rules

h) Time Restriction Rules

3.1.1.1 Application Rules:

These types of rules will govern whether some programs are allowed to access the Internet or not. The particular types of programs we will monitor are P2P sharing programs, instant messaging programs, trojans and other types of malicious software.

3.1.1.2 Protocol Rules

The administrator will select a list of protocols which will be allowed or denied. Consequently, some programs using certain protocols will not be allowed to access the Internet. Protocol rules are more broader than Application Rules in the sense that they can deny more than one application unlike application rules.

3.1.1.3 Port Rules

These rules will enable the administrator to block traffic from certain ports. In other words the programs using selected ports will not be able to access the Internet.

3.1.1.4 Bandwidth Rules

These rules will be used in restricting the real-time bandwidth usage of the user groups. It will be useful in that it will prevent unfair usage of the network resources.

3.1.1.5 Download Rules

Like bandwidth rules, the main purpose of download rules is to establish a fair share of network resources. This rule will restrict the user in terms of total amount of downloads.

3.1.1.6 File Download Operation Rules

We have three criteria for file downloads from the Internet. The first criteria is about the name of the file in question. We can allow or deny downloads of files according to the keywords in their name. The second criteria is the file extensions. The administrator will be able to control which file types are permitted for download. The last criteria is the size of the file. Administrator can create rules to determine maximum file size which can be downloaded.

3.1.1.7 URL/Domain/IP Access Rules:

We have three access rule criteria for general Internet access. The domains, URLs and IPs will be categorized into different sets of groups according to their content. Access to these groups will be allowed or denied according to the rules defined. Also, our policies will use these rules to give different access rights to different user groups.

3.1.1.8 Time Restriction Rules

This rule restricts the Internet usage in terms of time and domain classes. The administrator will be able to choose which domain classes will be restricted access between which time intervals and/or total amount of time.

3.2 SmartGuardian Crawler

Domain classes are an important part of the SmartGuardian Gateway. These classes comprises of domains with similar content. Domains will be inserted to the predefined classes. Web crawler will use data mining concepts to classify domains. The administrator will initiate the web crawler execution and he will be able to schedule crawling events. These domain classes will then be used to filter web sites according to the rules defined to user groups by the administrator.

There are 18 different groups incuding the uncategorized class which is used when a site cannot be classified to fit in any group. Other than this, the administrator will be able to create his own user defined categories and add his domains. Note that the web crawler will not add domains to the user defined groups.

The default groups are described next.

3.2.1 Domain Classes

We will have 18 groups classified for their content.

1. Pornography/Nudity:

These sites have content which may be offending to some users. They are adult sites.

2. Shopping:

These sites provide e-commerce services.

3. Social:

These sites include news sites, government sites, politics, religion and art related sites.

4. Criminal Activities:

These sites are mostly illegal in many countries, they include hate promoting sites, sites that describe how to prepare bombs or sites that advocate hacking.

5. Extreme:

These sites have violent content which is harmful to the development of children.

6. Games/Gambling:

These sites are one of the main factors of employee inefficiency. Employees consume both money and time at these types of sites.

7. Entertainment:

Includes web sites dealing with television shows, movies, animations, humor, fan clubs and so on.

8. Information/Communication:

This class includes sites that contain magazines, news and email servers.

9. IT:

This category is formed of web sites that are related with software, hardware, web hosting and IT services.

10. Drugs:

Sites in the drugs class contain information about illegal drugs, alcohol, and tobacco.

11. Sports:

These sites contain sports related information.

12. Human Resources:

These are career sites or company human resources pages.

13. Health:

These contain sites referring to medicine and medical care.

14. Finance/Investment:

These sites offer financial services. These include banking sites, online trading sites, investment sites and stock market sites etc.

15. Internet Risk/Fraud:

These include sites violating privacy, hacking sites, phishing sites that deceive surfers, and other malicious sites that risk security.

16. Education:

These include sites that are university web sites, general encyclopedia sites, research center sites which offer academic information.

17. Chat:

These are web sites that serve instant messaging services and chat servers.

18. Uncategorized:

Sites that do not fit into any other group are inserted into this group.

User Defined Categories:

The administrator will be able to create his own user defined categories and add his domains.

3.3 SmartGuardian Reporter

Reporting functionality will constitute an important part of our program. It will enable the administrator to take a snapshot of the network activities and to have a general view of the system. Reporter will be fairly flexible and the administrator will be able to control and manage the viewing properties (which info will be displayed). In terms of functionality, the following properties will be available.

3.3.1 Real Time Bandwidth Monitoring

The administrator will be able to view the total system bandwidth and user-wise bandwidth usage. This will give an idea about the network traffic load and thus the administrator can derive an optimal time for network maintenance operations and also for other operations the administrator plans to do. The advantage of viewing user-wise bandwidth usage is knowing the main consumers of the network resources and if it causes a bottleneck, necessary action will be taken appropriately.

The administrator will be able to change the logging time of bandwidth information, changing to suit the hard disk capacity and/or other limitations.

3.3.2 Domain and Time Statistics

This functionality of reporter will serve an important part in the system. The main aim in these kinds of programs are to have full monitoring over the user internet usage habits. The monitoring facility in our software will easily show which sites are accessed most frequently, and the time spent viewing them. All this information will be available for both user-wise and total usage. Thus, the company management will be able to compute how much time the employees spend with the work related sites and how much for their own pleasure.

3.3.3 Download Monitoring

According to download rules defined, some file requests may be turned down and other served files may be logged. The administrator will see which files are downloaded by which users. Again this will be logged as to giving the administrator an idea about which users request large files and/or files which are not permitted. Another thing to be taken into mind is that the administrator will monitor total and user-wise download sizes. This will be indicative in download amounts the users account for.

3.3.4 Unauthorized Activities

Most interesting information for the network administrator in terms of security will be unauthorized/denied access requests by the users. Denied requests of each user will be logged and shown to the administrator to warn him about potential dangers. For example, if a user tries to access to a hacking site, provided that this is unauthorized, it will be logged and reported extensively to the administrator.

3.4 SmartGuardian Controller

SmartGuardian Controller consists of a user friendly interface for administration of the whole system. Security is a crucial issue for these part of the software therefore all the transactions will be done through HTTPS protocol. All the user activities are kept in logs in order to audit the system in future. The functionality of this interface relies on the flexibility and stability of PHP and Java technologies.

SmartGuardian Controller greets the user with a login screen. After authentication, main menu will be displayed. In this menu, there will be links to wall settings, crawler settings, system settings and user settings.

Behavior of SmartGuardian Wall depends on the rules (which are mentioned in part 3.1.1) defined by users of the system. In Wall settings, user can define new rules, modify or delete existing ones.

In crawler settings, the administrator can manage the schedule of the crawler. The administrator can also create, delete, and modify his or her own domain classes (user defined classes).

System settings will contain general options which affect the whole system and options which do not fit into to any other part of the menu.

In user settings, the administrator can create, delete, and modify both users of the system and users of the network.