4.1 Cryptography: A New Symmetric Encryption Scheme

Cryptography, the mathematical science of secret communication, has a long and distinguished history of military and diplomatic uses dating back to ancient Greeks. Today, the ability to ensure the secrecy of military or diplomatic communication is as vital as ever. But cryptography is also becoming more and more important in everyday life, as the growth of computer networks for business transactions and communication of confidential information need safe, quick and inexpensive ways of exchanging information.

Encryption schemes are generally divided into symmetric and asymmetric key techniques. Symmetric encryption techniques are simple, but use the same key for both encryption and decryption, and hence needs a secure key transfer prior to real communication. Asymmetric encryption techniques do not need key transfer as they use different keys for encryption and decryption, but the encryption methods involve complicated mathematical operations making it unsuitable for encryption of large amounts of information. A combination of the two is attractive since asymmetric encryption technique can be used to transmit a secret key which is then used for encrypting the message using symmetric encryption scheme. The following is a symmetric encryption scheme based on the mathematical inhability to get exact solution from *n* simulations equations with more than *n* number of variables.

Figure 4.1.1 shows a typical symmetric encryption scheme involving Alice as sender, Bob as receiver and Eve as an adversary. Let us consider the following simultaneous equations.

$$K_{11}X_1 + K_{21}X_2 + K_{31}X_3 + K_{41}X_4 + K_{51}X_5 + K_{61}X_6 = P_1$$

$$K_{12}X_1 + K_{22}X_2 + K_{32}X_3 + K_{42}X_4 + K_{52}X_5 + K_{62}X_6 = P_2$$

$$K_{13}X_1 + K_{23}X_2 + K_{33}X_3 + K_{43}X_4 + K_{53}X_5 + K_{63}X_6 = P_3$$

$$K_{14}X_1 + K_{24}X_2 + K_{34}X_3 + K_{44}X_4 + K_{54}X_5 + K_{64}X_6 = P_4$$

Let $K = \{K_{11}, \ldots, K_{64}\}$ be the keys for encryption and $P = \{P_1, \ldots, P_4\}$ be the plain text message to be sent. To encrypt the message P, Alice will pick randomly one of the infinite possible $X = \{X_1, \ldots, X_6\}$ values from the above simultaneous equations and send that to Bob. For Bob, it is a simple mathematical calculation to get back *P* as he knows *K*.

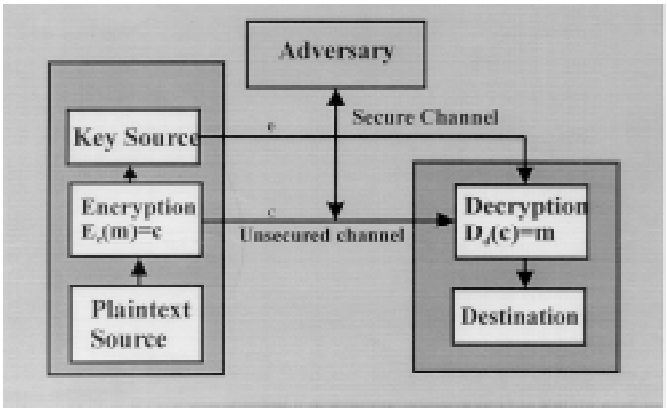
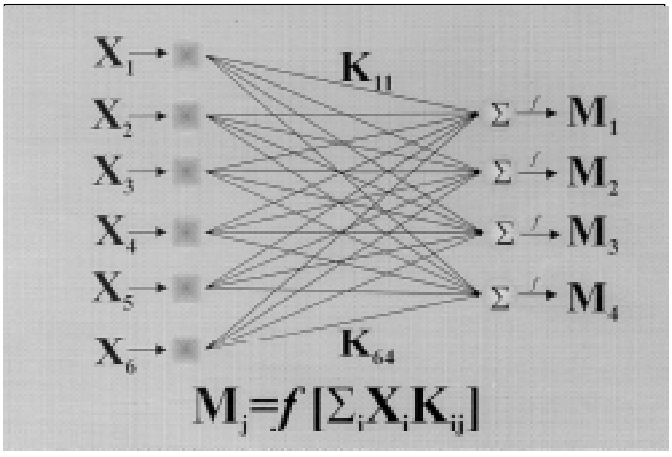


Fig. 4.1.1. Block diagram of symmetric eneryption scheme.

Eve, by knowing only *X*, can make exhaustive key search, but will fail due to limited computing power available in reality. This new technique is protected from possible attack by matching the frequency of frequently used words like 'the', 'is', 'are', etc. as the encrypted message *X* is different for the same plain text message *P* at different instances of time because it has been randomly picked. The security can be further enhanced by using different keys for each set of plain text *P*, which is derived from previous keys in a mutually agreed way. This prevents Eve from breaking the code by predicting a number of possible common words related to the source of the message, i.e "ocean", atmosphere", "earth", etc. for messages from C-MMACS.

Picking random *X* is the key issue in this kind of encryption scheme since, otherwise, the system may become ill-conditioned. This can be achieved in many ways

Fig. 4.1.2. 6-input, 4-output neural network model for representing 4 simultaneous equations in 6 unknown variables.



and a method using back-propagation algorithm is discussed here. Figure 4.1.2 shows a six-input, four-output network, which represents the above set of equations. A random input pattern is obtained by randomly choosing the initial input values of the network and training the network for the desired output pattern using the iterative algorithm,

$$\Delta X_i(n) = \alpha \Delta X_i(n-1) + \eta \sum_j K_{ij} e_j(n) f^{-1}(V_j(n))$$

where *i* and *j* are the input and output nodes respectively, η the learning rate and α the momentum parameter. $e_j(n)$ is the error energy for output node *j* so that $e_j = M_j - Y_j$ where $Y_j = f(V_j)$ and $M_j = f(P_j)$; f is a non linear function and $V_j = \Sigma_i K_{ij} X_i$. The disadvantage of this kind of method is the time and resources required to encrypt a big message. Work is in progress to develop a better algorithm for randomly picking a solution.

(G K Patra)

4.2 Modelling in Epidemiology

The National Institute of Epidemiology initiated a collaboration with C-MMACS to understand SIMLEP, a model to simulate the spread and control of leprosy. We carried out a detailed sensitivity analysis of SIMLEP. It was found that the model is insensitive to the age group of the individuals getting infected. On the other hand, it was found that the model is most sensitive to the incidence rate followed by the force of infection and the transmission rate. The other parameters which were studied in detail were the proportion of population with natural immunity, proportion with short incubation period amongst newly infected cases who will not develop strongly contagious clinical disease and, the proportion with short incubation period amongst newly infected cases who will develop strongly contagious clinical disease; it was also checked whether these parameters have any relationship with either the incidence rate or the force of infection or the transmission parameter.

The model has been tuned for the Indian situation by considering the parameters important to the Indian population and to the regions considered in India for the leprosy programme. In order to make SIMLEP run for Indian situations and use that as a tool to control the leprosy transmission in India, further work is needed as the model is densely connected with many parameters. The work is in progress.

(N K Indira)

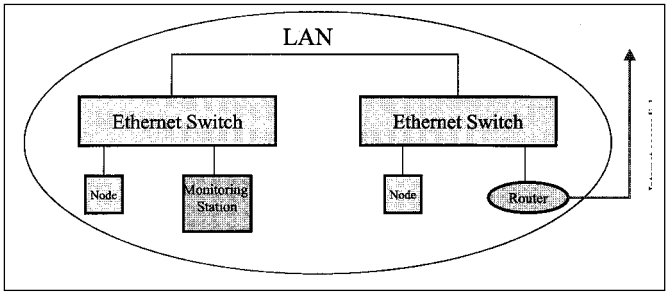
4.3 Measurement of Bandwidth Utilisation of Internet Access Link Using Simple Network Management Protocol

The recent explosive growth of Internet, in terms of its size, number of applications, amount of information exchanged etc. made it clear that traditional approaches to measure and control the usage of bandwidth is no longer adequate in avoiding congestion on the Internet. One of the recent surveys reveals that over 15,000 web sites are added to the net everyday. This translates to a new website in every 6 seconds and 100 Gigabytes of data in every hour. In order to ensure a fair and decent availability of bandwidth to genuine Internet users, sophisticated techniques will have to be developed to monitor and control Internet bandwidth.

One of the most essential and crucial components of a bandwidth manager is an accurate and efficient technique to measure the bandwidth utilisation in real time which many commercially available managers do not have. Since the control signal generated to regulate the bandwidth in bandwidth managers is dynamically calculated as a function of the actual bandwidth utilisation at that time, an error in the measurement causes an error in control signal and this leads to an unbalanced bandwidth management.

We have designed and developed a system to measure the bandwidth utilisation of Internet access link in real time. Our approach is based on Simple Network Management Protocol (SNMP), a well known member of TCP/IP protocol family. Just like any other network application, our approach also follows the client server architecture in which a machine, called monitor, in the Local Area Network (LAN), will periodically probe the Internet access router, for both incoming and outgoing traffic statistics. The access router running standard SNMP daemon acts as the server and the monitor which probes and collects the traffic counts from the router acts as the client. Figure 4.3.1 shows the client server architecture consisting of the monitor as the client and the Internet router as the server in LAN

Fig. 4.3.1. Physical deployment of the monitor and access router in the LAN environment.



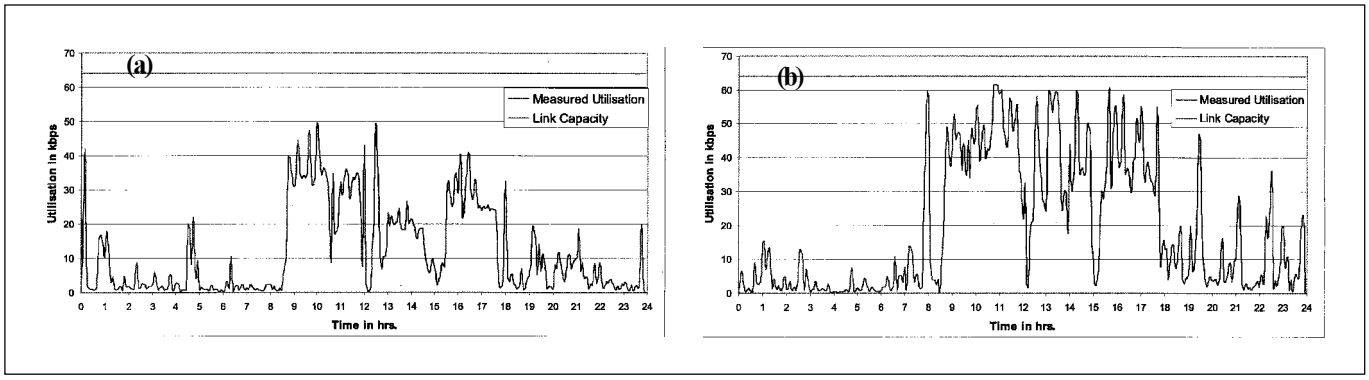


Fig. 4.3.2. Incoming (a) and outgoing (b) traffic measured on C-MMACS 64 kbps WAN link on a regular working day.

environment.

Two variables, namely *IfInoctets* and *IfOutoctets*, from the interface class of the standard Management Information Base (MIB) are used to send the query to the router. The variables used are standard variables and all implementations of SNMP support these variables. So the approach can work well without any modification on the server side.

The formulae used to calculated the bandwidth utilisation are:

$$InputTraffic = \frac{|IfInOctets(t2) - IfInOctets(t1)| \times 8}{|t2 - t1| \times 60 \times 1024}$$

$$OutputTraffic = \frac{|IfOutOctets(t2) - IfOutOctets(t1)| \times 8}{|t2 - t1| \times 60 \times 1024}$$

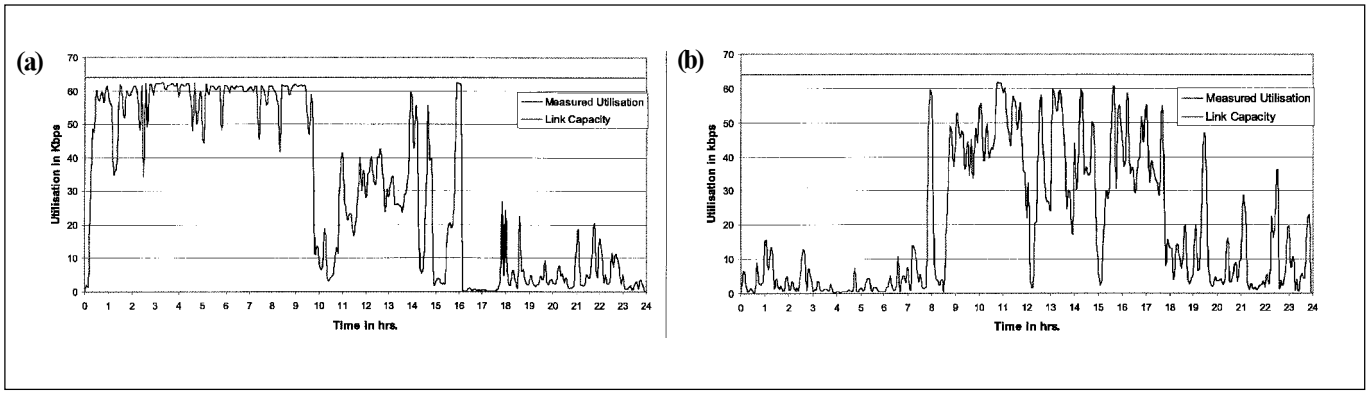
where *(t2-t1)* is the time difference in minutes between two queries, *IfInOctets(t2) - ifInOctets(t1)* is the difference in input traffic count during this period *(t2-t1)* in bytes and *IfOutOctets(t2) - ifOutOctets(t1)* is the difference in output traffic count during the same period in bytes.

By reducing the value of *(t2-t1)* appropriately, the accuracy of measurement can be improved. In our experiment we have chosen *(t2-t1)* as 2 minutes, which was adequate to produce an accurate and continuous graph of utilisation verses time over a period of 24 hrs.

From the software implementation point of view, the system consists of two independent modules. The first module, called the data acquisition module, is designed to run round-the-clock and obtains input and output traffic statistics from the router. The second module is the GUI, which plots the incoming and outgoing traffic against time on a linear scale.

Figures 4.3.2 *a* and *b* show the measured input and output traffic on our 64 kbps link on a normal working day. It is evident that both incoming and outgoing traffic are minimum during after office hours and is high during the office hours. Figures 4.3.3 *a* and *b* give the input and output traffic when a file transfer protocol (ftp) is initiated to transfer few huge files with sizes of the order of 250 megabytes from Internet to C-MMACS LAN. The file transfer was started at 0:10 hrs and was allowed to continue till 9:15 a.m. so that it does not affect the data transfer initiated by regular users during office hours. It is interesting to observe that the file transfer keeps only the incoming channel busy and outgoing channel is almost free during this period. The area under the graph between 0:10 hrs and

Fig. 4.3.3. Incoming (a) and outgoing (b) traffic during a ftp session.



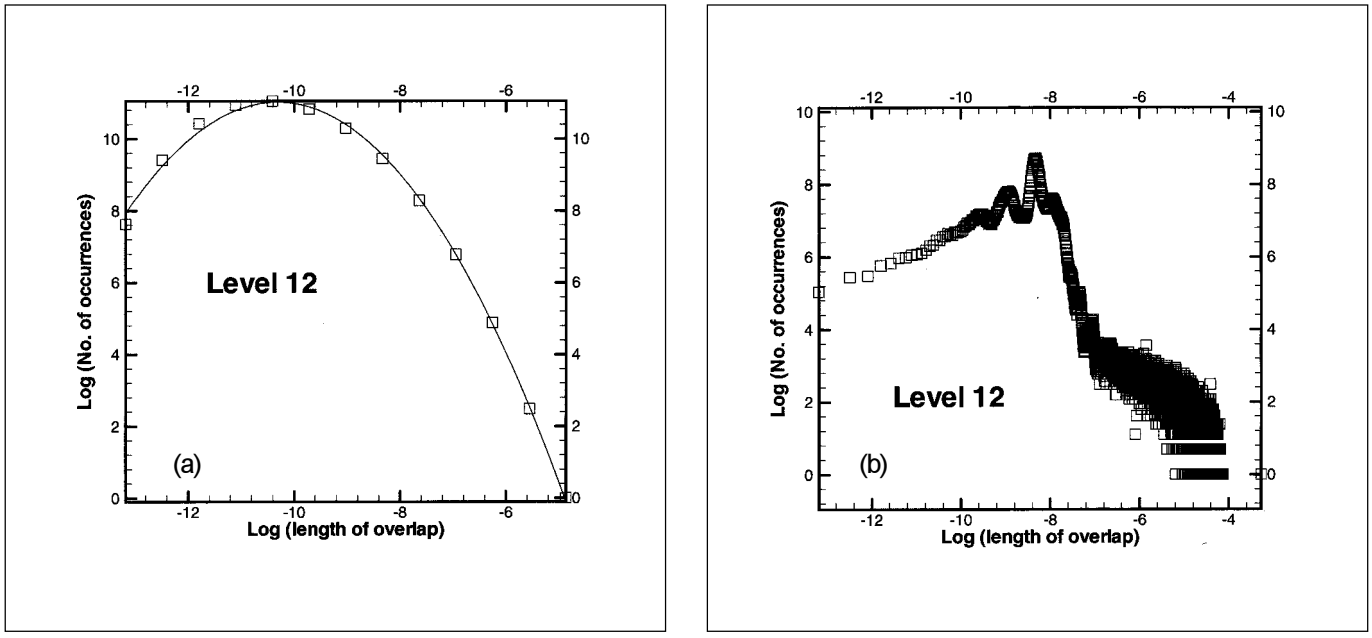


Fig. 4.4.1. Log-Log plots of contact areas versus overlap for two cantor sets as one is moved over another in elementary steps; (a) is without noise and (b) is with 20% noise. The Cantor sets have been taken in the twelfth level of construction.

9:15 hrs gives the total number of bits transferred to C-MMACS LAN during the ftp session.

Some of the advantages of the SNMP based approach over traditional approaches to monitor the bandwidth utilisation are: the SNMP based approach does not demand the monitoring station to be physically positioned between the access router and the LAN or in the same LAN segment where the router is positioned. The monitoring station can be physically positioned anywhere on the local network or on the Internet. The failure or performance degradation of the monitoring station will not affect the performance of the link except that monitoring is not possible during the failure of the monitoring station. Also no dedicated machine is required for the purpose of monitoring since the amount of computation and processing required are negligibly small compared to the power of a standard personal computer. From the security point of view, since the monitoring station does not require any privilege on the access router except a read permission, the approach does not generate any security loopholes.

It may also be pointed out that, from the graphs, in addition to estimating the time during which the access link is down, it is also possible to distinguish the failure of the link from the failure of the data acquisition module. A discontinuity in our graph represents the failure of the data acquisition whereas the graph with zero value represents the link failure.

(V Anil Kumar, G Parthasarathy, R Vanitha*; *Periyar University, Salem)*

4.4 Distribution of Contact Areas Between Rough Surfaces

The dynamics of two rough surfaces in motion past one another while in contact is a problem with applications in several areas. In particular, the problem has relevence in the area of plate tetonics where earthquakes are caused by the motion at the boundaries of plates. We have developed algorithms to compute the distribution of contact areas of two surfaces and applied it to the case of two fractal surfaces in contact. Specifically, the case of two cantor sets in contact was studied and the results are as shown in Fig. 4.4.1 *a*. Note that the symmetry of the cantor set is reflected in the distribution. In Fig. 4.4.1 *b*, we show the results when noise (20%) has been added to the system. Further investigations are planned with a variety of surfaces.

(T R Krishna Mohan)