A Covariance Analysis Model for DDoS Attack Detection*

Shuyuan Jin
Department of Computing
HongKong Polytechnic University
HongKong, China
cssyjin@comp.polyu.edu.hk

Daniel S. Yeung
Department of Computing
HongKong Polytechnic University
HongKong, China
csdaniel@inet.polyu.edu.hk

Abstract—**This paper discusses the effects of multivariate correlation analysis on the DDoS detection and proposes an example, a covariance analysis model for detecting SYN flooding attacks. The simulation results show that this method is highly accurate in detecting malicious network traffic in DDoS attacks of different intensities. This method can effectively differentiate between normal and attack traffic. Indeed, this method can detect even very subtle attacks only slightly different from normal behaviors. The linear complexity of the method makes its real time detection practical. The covariance model in this paper to some extent verifies the effectiveness of multivariate correlation analysis for DDoS detection. Some open issues still exist in this model for further research.**

I. Introduction

Distributed Denial of Service (DDoS) attacks, which aim at overwhelming a target server with an immense volume of useless traffic from distributed and coordinated attack sources, are a major threat to the stability of the Internet [1]. But, for four reasons, it is difficult to detect an ongoing DDoS attack: Firstly, because a DDoS attack has to be detected on-line, there is little time to detect and confirm an ongoing DDoS attack. Normally the system administrators or security experts have to ascertain the attacks or trace back the attackers in less than one hour. Secondly, some Internet worms' propagation may also directly result in DDoS, which makes DDoS detection much more complex. Thirdly, normal defense measures such as rate-limiting, packet-filtering, tweaking software parameters or equipping more servers are all useful but limited in their capabilities. Finally, the exact distinction between DDoS attacks and flash crowds remains an open issue. [2] has proposed a taxonomy, but it is based only on HTTP protocol and web events.

It is essential that we be able to detect DDoS attacks fast, accurately and in real time. DDoS attacks exhaust host resources or the network bandwidth. It is consequently important to detect resource usage changes and reduce the detection time. Such abnormal changes could be detected statistically. For example, the entropy method in [3] uses frequency-sorted distributions of selected packet attributes in a time window to compute entropy and use the entropy changes to indicate the anomalies. In [4] the Chi-Square statistics approach is used to identify the anomalies. [5] Demonstrates how Aderson-Darling statistical method is used to detect network traffic changes. The adaptive sequential and batch-sequential methods in [6] employ statistical analysis of data from multiple layers of network protocols to detect very subtle traffic changes.

The problem with statistics-based detection methods is that it is not possible to determine with certainty the normal network packet distribution. Rather, it can only be simulated as a uniform distribution. Some papers suggest the usage of clustering methodology to formulate the normal patterns. For example, in [7], hop-count information inferred from Time To Live (TTL) value in the IP header is used to cluster the normal address prefixes. [2] uses a network aware clustering method to distinguish between the HTTP-based flash events and DDoS attacks. One of the advantages of clustering methods, over statistics-based methods, is that they do not rely on any prior known data distribution. There exist many variables that can be used to identify normal network patterns: client characteristics such as the source IP address or Round-Trip Time (RTT), network traffic patterns such as the different protocol distribution, packet rate and flow interval, and server patterns such as the number of clients and their distribution, per-client request rate and request file distributions. The choice of which variables selected as an adequate clustering criterion remains an open problem. Modeling detection problems into control theory [8] and material flow [9] control problems may also provide novel perspectives on the DDoS detection.

The covariance analysis method proposed in this paper is also a statistics-based method, but unlike the works in [4] [5] and [6], it does not rely upon any presumptions on the normal network packets distributions. This method is effective because it accurately detects DDoS attacks of different intensities, and because the method is simple, it can be implemented on-line.

This paper claims three main contributions:

- This paper aims at discussing the effect of multivariate correlation analysis on the detection of DDoS attacks. Although in the experiments we only use all flags in the control field of TCP header as raw data and only use covariance matrices to detect SYN flooding attacks, it may not obstruct the importance and generalization of our proposed approach. In [10], it was shown that SYN flooding attacks could also be detected using the non-parametric Cumulative Sum (CUSUM) method. The difference between these two

* This research work is supported by a Hong Kong RGC project research grant number B-Q571

0-7803-8533-0/04/$20.00 (c) 2004 IEEE

papers lies in that we aim at proposing a generic correlation method to detect DDoS attacks, and the intention of experiments of detecting SYN flooding attacks only seek to exemplify this idea.

- This attempt will be helpful in the research of feature selection for detecting DDoS attacks. In fact, many parameters were used in detecting DDoS, such as source IP address and RTT [2, 7 and 10]. But no systematic way of choosing these parameters has been proposed. [11] presents a feature ranking and selection for IDS and [16] discusses a novel feature reduction technique. Their experiments are based on datasets described in [12]. They did not discuss some other parameters such as flow inter-arrival time or packet rate. Nor did it consider the effect of any correlation coefficients on detection.
- This paper proposes an effective, on-line detection method based on covariance analysis. The method could not only differentiate between the normal and attack traffic in flooding situations, but also detect the subtle attacks that expose few apparent differences from normal behaviors. It is well known that "The most difficult part for defending against DDoS attacks is that it is very hard to differentiate between normal traffic and attack traffic. This is the fundamental problems of the Internet [15]". Besides, the linear complexity of our method makes DDoS real-time detection efficient and practical.

The rest of this paper is organized as follows: Section II describes the covariance analysis detection methodology used in DDoS detection. Section III validates the covariance detection method in simulations and experiments of detecting SYN flooding attacks. Section IV analyzes the method's effectiveness. Finally, conclusions and discussions are proposed in Section V.

II. Overview of covariance analysis method

The main idea is that the characteristics of an information system could be described by the correlations among its features. Ranking and selecting the known parameters as features could help to identify the patterns of the Internet, while the correlations among the features may provide additional essential information. The correlations are expected to be sensitive enough to flag some changes. In terms of correlation, the normal patterns will be different from the abnormal patterns. In this sense, the correlation may be selected as a change indicator, and any changes or abnormal activities will definitely change the correlation coefficients of these features gotten in the normal situations. So detecting the correlation changes could determine the occurrence of the change. In statistics theory, the effectiveness of this change detection method is obvious, and its efficiency is determined by the suitable data volume we could gather in a limited observation window. The details of the method are described as below:

Assume there are p features, $f_1,...,f_p$, which compose a random vector $X=(f_1,...,f_p)'$. Let $x_1,...,x_n$ are the n observed vectors, $x_i=(f_1^i,...,f_p^i)$ is the i^{th} observed vectors. $f_i^{l,j}$ is the value of f_i in the j^{th} observation during the l^{th} time interval T_l. We define a new variable y and the covariance matrix M to characterize the variable y as follows:

$$y_l=\begin{pmatrix} f_1^{l,1} \dots f_p^{l,1} \\ f_1^{l,2} \dots f_p^{l,2} \\ \vdots \ \ddots \ \vdots \\ f_1^{l,n} \dots f_p^{l,n} \end{pmatrix} \tag{1}$$

$$M_{y_l}=\begin{pmatrix} \sigma_{f_1^l f_1^l} & \sigma_{f_1^l f_2^l} & \dots & \sigma_{f_1^l f_p^l} \\ \sigma_{f_2^l f_1^l} & \sigma_{f_2^l f_2^l} & \dots & \sigma_{f_2^l f_p^l} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{f_p^l f_1^l} & \sigma_{f_p^l f_2^l} & \dots & \sigma_{f_p^l f_p^l} \end{pmatrix} \tag{2}$$

We define variable z_l as the distance between two matrices M_{y_l} and the mean of M_{y_l} or $E(M_{y_l})$:

$$z_l=\left\|M_{y_l}-E(M_{y_l})\right\| \tag{3}$$

where z_l measures the change, or the anomaly. To simplify the problem description, the following distance function of two matrices is used:

$$\|M_1-M_2\|=\sqrt{\sum_{1\le i,j\le p}(a_{i,j}-b_{i,j})^2},\ \forall a_{i,j}\in M_1,\ \forall b_{i,j}\in M_2,\ 1\le i,j\le p \tag{4}$$

In the normal situation, one can find a point c and a constant a for z_l in equation (3), which satisfy $|z_l-c|<a,\ \forall l\in Z$. The constant a is selected as the upper threshold of the i.i.d $|z_l-c|$. So for the observed data gathered during the l^{th} time interval, we calculate the corresponding z_l, if $|z_l-c|>a$, the abnormal behaviors could be determined.

III. SYN Flooding simulations and detection

We select all the flags in control field of TCP header as features in the covariance model. Each flag occupies only 1 bit in the TCP header. As in [13], a brief description of each bit is given in the following table:

Table 1 Description of flags in the control field of TCP header

Flag	Description
URG	The value of urgent pointer field is valid
ACK	The value of acknowledgement field is valid
PSH	Push the data
RST	The connection must be reset
SYN	Synchronize sequence numbers during connection
FIN	Terminate the connection

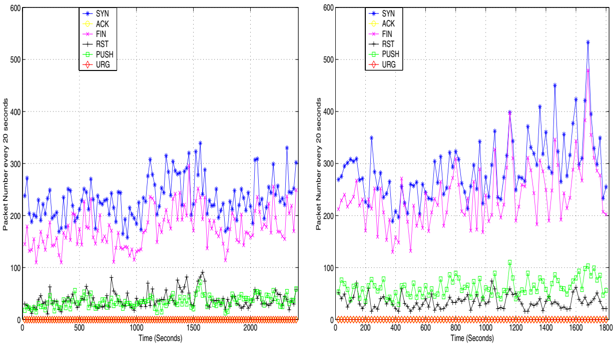
As we know, in SYN flooding attacks, the numbers of SYN and FIN do not match. Our method tries to use the covariance of each pair of the above six flags to detect the SYN flooding attacks. The dataset we use are described in table 2. Each of these two traces contains an hour's worth of all wide-area traffic between Digital Equipment Corporation and the rest of the world. The traces were gathered at Digital's primary

Internet access point, which is an Ethernet DMZ network operated by Digital's Palo Alto research groups.

TABLE 2 THE TCP TRACE DESCRIPTION

Trace	Start Time	TCP Packets
dec-pkt-1	22:00, Wed March 8th, 1995	3.3 million
dec-pkt-2	02:00, Thu March 9th, 1995	3.9 million

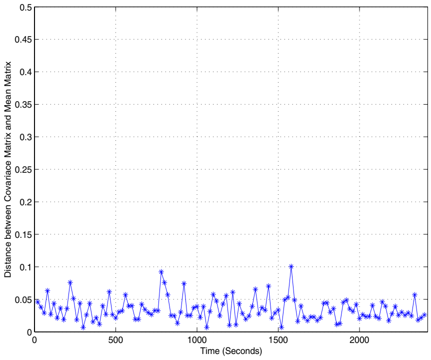
In the simulation, these two traces represent normal traffic. We parse the two traces and extract the control field value of each packet in the traces. We select 20 seconds as a time interval. Figures 1 (a) and (b) represent the characteristics of different kinds of packet number under normal operations. These two figures give a general description of two traces in the selected time interval in terms of different packet types.



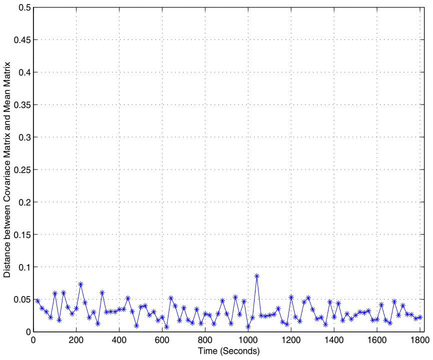
(a) dec-pkt-1 (b) dec-pkt-2

Figure 1 Normal Packet Number according to different flags in 20 second interval

According to equation (2), we calculate all matrices *M*s in 20 seconds in the trace dec-pkt-1 and dec-pkt-2 separately. Then we find the corresponding z_ls defined by equation (3), according to the distance definition given in (4). The result is described in figures 2 (a) and (b), which illustrate the distance between the covariance matrix and the mean of all covariance matrices under normal situations.



(a) dec-pkt-1

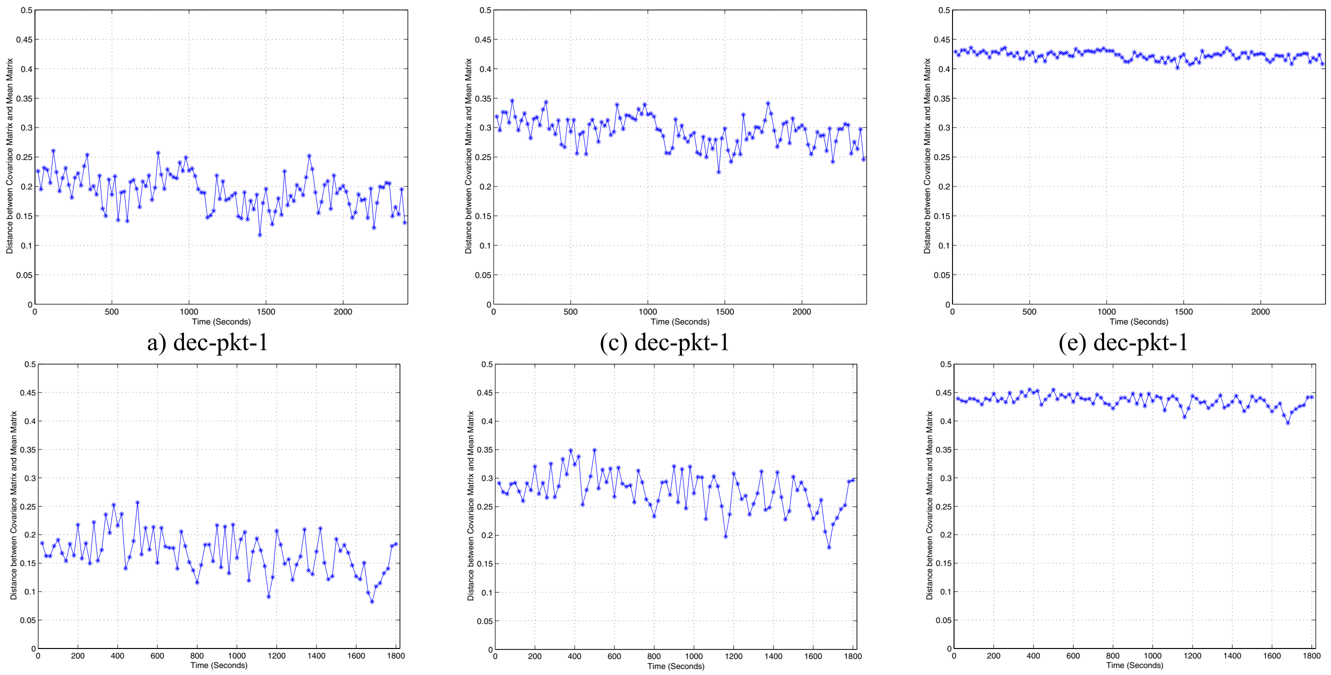


(b) dec-pkt-2

Figure 2 Covariance Matrix Distance under Normal Operations

The abnormal traffic is simulated the same way as described in [10]. In a DDoS attack, the severe effect on the victim comes from the aggregate flooding rate, rather than from only one single attack source. So the sensitivity needs for attacker-nearest-router located and victim-nearest-router located DDoS detection are different. The attacker-nearest-router located DDoS detection focuses on preventing the output of internal malicious traffic to the Internet, which corresponds to the low attack packet rate detection. Victim-nearest-router located DDoS detection mainly focuses on protecting the victims, which corresponds to the high attack packet rate detection. To attack a protected server, the aggregate flooding packet rate should be larger than 14000 [10]. In order to detect our method’s sensitivity in both aspects, we simulate two situations: 500 SYN packets per second to the victim-nearest-router located detection, and 35 packets per second to the attacker-nearest-router located detection. Each is the minimal attack packet rate in each case.

In order to show the trends of the effects of different attack packet rates on the covariance matrices distance between the normal and the abnormal one, we also simulate the situation of an attack rate of 80 packets per second. From figure 3, we can see that this covariance matrix distance method performs well in detecting not only large flooding DDoS attacks, but also subtle DDoS attacks, even if the attack rate is 35 packets every second (figures 3 (a) and (b)) where at least 400 stub networks with this attack rate would be needed to make a successful DDoS attack. So the sensitivity of this method is obviously high. It will have a 100% detection rate if the threshold of distance is set as 0.1. The results showed that under very subtle attacks, the method could still detect them even if they exposed very similar behaviors to the normal ones. The results also showed the detection accuracy trends when attack rate increased. It is clear that increasing the attack rate will greatly enlarge the gap between the distance under normal operations (in figure 2) and that under abnormal operations (in figure 3).



a) dec-pkt-1 (c) dec-pkt-1 (e) dec-pkt-1

(b) dec-pkt-2 (d) dec-pkt-2 (f) dec-pkt-2

Figure 3 Sensitivity of Covariance Matrix Distance Method under Different Attack Rates. (a)& (b), (c) & (d), and (e) & (f) show the attack rates of 35, 80 and 500 SYN packets per second, respectively.

IV. Method Effectiveness Analysis

The effectiveness of the proposed method is shown by the algorithm's complexity analysis.

Assume in the time interval T_l, n vectors are observed. For each feature we obtain n values. For example, $\forall i, \forall j, 1 \le i, j \le p$, we get $f_i^l = (f_i^{l,1}, f_i^{l,2}, ..., f_i^{l,n})$ and $f_j^l = (f_j^{l,1}, f_j^{l,2}, ..., f_j^{l,n})$. Then the covariance coefficient $\gamma(f_i^l, f_j^l)$ of f_i and f_j could be represented as:

$$\gamma(f_i^l, f_j^l) = \frac{COV(f_i^l, f_j^l)}{\sqrt{D(f_i^l)D(f_j^l)}} \tag{5}$$

where

$$COV(f_i^l, f_j^l) = E(f_i^l f_j^l) - E(f_i^l)E(f_j^l)$$
$$D(f_m^l) = E(f_m^l)^2 - (E(f_m^l))^2, m = i, j$$

Since $\forall k, i,\ 1 \le k \le n,\ 1 \le i \le p,\ f_i^{l,k} \in \{0,1\}$, we have $(f_i^{l,k})^2 = f_i^{l,k} \in \{0,1\}$. So the calculation of $\gamma(f_i^l, f_j^l)$ could be simplified as:

$$\gamma(f_i^l, f_j^l) = \frac{COV(f_i^l, f_j^l)}{\sqrt{D(f_i^l)D(f_j^l)}} = \frac{m/n - PQ}{\sqrt{(P - P^2)(Q - Q^2)}} \tag{6}$$

where

$$P = E(f_i^l) = \frac{1}{n}\sum_{k=1}^{n} f_i^{l,k} \sim O(n)$$
$$Q = E(f_i^l) = \frac{1}{n}\sum_{k=1}^{n} f_j^{l,k} \sim O(n),$$

And m is the occurrence number of $f_i^{l,k} = f_j^{l,k} = 1$ gotten by comparison operations.

The value of each entry in the covariance matrix could be calculated simply by the summation and comparison operations. The complexity of the method is therefore linear which makes the on-line detection practical

V. Conclusions and Discussions

This paper discusses the effects of multivariate correlation analysis on the detection of DDoS attacks. In terms of correlation, the normal patterns will be different from the abnormal patterns. In this sense detecting the correlation changes among different features could determine the occurrence of the anomalies. A two variables covariance model is presented in this paper as a possible approach to detecting the DDoS attacks. One could explore the three or more variables correlation model in future research.

Compared with existing statistical methods used in DDoS detection [4, 5 and 6], the covariance analysis method proposed in this paper offers the advantage of independence from packet distribution assumption.

In this paper, all flags in the control field of the TCP header are used as features in the covariance analysis model. The simulation results showed that this method is highly accurate in detecting SYN flooding attacks in DDoS. This method could also effectively differentiate between normal and attack traffic. What is more, it could detect attacks of very subtle intensity, exposing near-to-normal behaviors. The results of high detection accuracy in the experiments and real-time effectiveness analysis reveal some impacts of the multivariate correlation analysis on the detection of the DDoS attacks.

However, our present method has three major limitations. First, there is no guarantee that the 6 flags are valid or sufficient features used in the covariance analysis model for DDoS detection. Second, no theoretical justification is provided for the high detection rate as demonstrated in the experiments. Third, how to select the appropriate observed time interval is still an open problem.

References

[1] Emergency Response Team. Result of the Distributed-Systems Inruder Tools Workshop. http://www.cert.org/reports/dsit_workshop-final.html, November 1999.

[2] J. Jung, B. Krishnamurthy, M. Rabinovich. Flash Crowds and Denial of Service Attacks: Characterization and Implications for CDNs and Web Sites. The Eleventh International World Wide Web Conference, Honolulu, Hawaii, May 2002.

[3] L. Feinstein, D. Schnackenberg. DDoS Tolerant Network. Proceedings of the DARPA Information Survivability Conference and Expostion(DISCEX'03), April 2003.

[4] L. Feinstein, D. Schnackenberg. Statistical Approaches to DDoS Attack Detection and Response. Proceedings of the DARPA Information Survivability Conference and Expostion(DISCEX'03), April 2003.

[5] C. Manikopoulos, S. Papavassiliou. Network Intrusion and Fault Detection: A Statistical Anomaly Approach. IEEE Communications Magazine, October 2002.

[6] R. B. Blazek, H. Kim, B. Rozovskii, A. Tartakovsky. A Novel Approach to Detection of Denial-of-Service Attacks Via Adaptive Sequential and Batch-Sequential Change-Point Detection Methods. Workshop on Statistical and Machine Learning Techniques in Computer Intrusion Detection, June 2002.

[7] C. Jin, H. Wang, K. G. Shin. Hop-Count Filtering: An Effective Defense Against Spoofed Traffic. ACM Conference on Computer and Communications Security (CCS'2003), October 2003.

[8] Y. Xiong, S. Liu, P. Sun. On the Defense of the Distributed Denial of Service Attacks: An On-Off Feedback Control Approach. IEEE Transactions on System, Man and Cybernetics---Part A: System and Humans, July 2001.

[9] J. Kong, M. Mirza, J. Shu, C. Yoedhana, M. Gerla, S. Lu. Random Flow Network Modeling and Simulations for DDoS Attack Mitigation. IEEE 2003 International Conference on Communications, May 2003.

[10] H. Wang, D. Zhang, K. G. Shin. Detecting SYN Flooding Attacks. The Twenty-First Annual Joint Conference of IEEE Computer and Communications Societies, INFOCOM, May 2002.

[11] S. Mukkamala, A. H. Sung. Feature Ranking and Selection for Intrusion Detection using Support Vector Machines. Presentations in Workshop on Statistical and Machine Learning Techniques in Computer Intrusion Detection, June 2002.

[12] KDD Cup 1999 Data for Intrusion Detection. http://kdd.ics.uci.edu/databases/kddcup99/kddcup.names. UCI Knowledge Discovery in Databases Archive.

[13] B. A. Forouzan. TCP/IP Protocol Suite, Second Edition. McGraw Hill, 2003.

[14] P. Porra, A. Valdes. Live Traffic Analysis of TCP/IP gateways. In the proceeding of ISOC symp. on Network and Distributed System Security(NDSS'98), March 1998.

[15] T. Peng, C.Leckie, L.Ramamohanarao. Protection from Distributed Denial of Service Attacks Using History-based IP Filtering. IEEE 2003 International Conference on Communications, May 2003.

[16] W. W. Y. Ng, R. K.C. Chang, D.S. Yeung, Dimensionality Reduction for Denial of Service Detection Problems Using RBFNN Output Sensitivity, to appear in Proceedings of International Conference on Machine Learning and Cybernetics, Xian, China, November 2003

Appendix

In equation (2) in Section II, the definitions of variable σ and variable μ are as follows:

$$\mu_{f_i^l} = E(f_i^l) = \frac{1}{n}\sum_{k=1}^{n} f_i^{l,k}$$

$$\sigma_{f_i^l f_j^l} = \frac{1}{n}\sum_{k=1}^{n}\left(f_i^{l,k} - \mu_{f_i^l}\right)\left(f_j^{l,k} - \mu_{f_j^l}\right)$$

where $1 \le i, j \le p$, p is the number of features. $1 \le k \le n$ is the number of observations during T_l, and $l,$ $1 \le l \le n$,is the number of time intervals.

In Section IV, equation (5) can be reduced to equation (6).

Since $\forall k, i,\ 1 \le k \le n,\ 1 \le i \le p,\ f_i^{l,k} \in \{0,1\}$, we have $(f_i^{l,k})^2 = f_i^{l,k} \in \{0,1\}$

$$E(f_i^l)^2 = \frac{1}{n}\sum_{k=1}^{n}(f_i^{l,k})^2 = \frac{1}{n}\sum_{k=1}^{n} f_i^{l,k} = E(f_i^l) \sim O(n)$$

$$D(f_i^l) = E(f_i^l)^2 - (E(f_i^l))^2 = E(f_i^l) - (E(f_i^l))^2 \sim O(n)$$

Clearly,

$$D(f_j^l) \sim O(n)$$

$$E(f_i^l f_j^l) = \frac{1}{n}\sum_{k=1}^{n} f_i^{l,k} f_j^{l,k} = \frac{m}{n} \sim O(n)$$

Here m is the occurrence number of $f_i^{l,k} = f_j^{l,k} = 1$. The value of the variable m could be obtained by simple comparison operations. No calculations are needed when $f_j^{l,k}$ or $f_i^{l,k}$ is equal to 0. Only when the two variables are equal to 1, m will be increased by 1.