IDS ログから算出される 情報エントロピー値の変動に注目した異常検出

竹森 敬祐† 三宅 優† 田中 俊昭† 笹瀬 巌‡

† （株）KDDI 研究所 〒356-8502 埼玉県上福岡市大原 2-1-15

‡ 慶應義塾大学理工学部 〒223-8522 神奈川県横浜市港北区日吉 3-14-1

E-mail: †{ke-takemori, yu-miyake, tl-tanaka}@kddi.com, ‡ sasase@sasase.ics.keio.ac.jp

あらまし 昨今，急速にインターネット全域に拡大する攻撃が深刻な問題になっている．ネットワークサービスに影響を与える大規模攻撃やワーム感染ホストが与える局所的な攻撃の確実な検出と迅速な対応, 被害規模の把握は重要な課題である. 本研究では, 侵入検知システム（IDS: Intrusion Detection System）のイベント出力に関する情報理論的な曖昧度を情報エントロピーによって算出し，その長期間の統計分布の平均と標準偏差を用いて，短期間の異常性を評価する手法を提案する．実際の IDS ログを用いて局所的攻撃の検出率に関する評価を行い，本手法が未検出率ならびに誤検出率を低減できること，従来からのイベント頻度を用いた異常検出手法と組み合わせることで確実に検出できることを示す．本手法をインターネットの攻撃概況指標へ適用することで，セキュリティ監視者の迅速な対応と情報交換に寄与する．

キーワード IDS ログ，情報エントロピー，異常検出，統計分布

An Anomaly Detection Technique for IDS Events using Deviations of Information Entropy

Keisuke TAKEMORI† Yutaka MIYAKE† Toshiaki TANAKA† and Iwao SASASE‡

† KDDI R&D Laboratories Inc. 2-1-15 Ohara, Kamifukuoka-shi, Saitama, 356-8502 Japan

‡ Dept. of Info.& Computer Science, Keio Univ. 3-14-1 Hiyoshi, Kohoku-ku, Yokohama-shi, 223-8522 Japan

E-mail: †{ke-takemori, yu-miyake, tl-tanaka}@kddi.com, ‡ sasase@sasase.ics.keio.ac.jp

Abstract Recently, rapid increasing attacks that influence network services have become critical issues on the Internet. A detection technique for large scale attacks and worm infected hosts needs to estimate degree of its propagation. In this research, we propose an anomaly detection technique for IDS (Intrusion Detection System) events using the information entropy. And the information entropy is adapted to a profiling approach which compares the current information entropy with mean and standard deviations of the past information entropies. Experimental evaluations with real IDS events show that the detection ratio of false positives and false negatives for the large scale attacks or the worm attacks on our approach is better than that using event counts on previous approach. Furthermore, the combination system of our approach and previous approach is able to detect potential issues perfectly. We also adapt the techniques to a thereat indicator, and its objective alarms effect with quick and reliable response for security operators.

Keyword IDS Log, Information Entropy, Anomaly Detection, Statistical Distribution

1. はじめに

近年，サイバーテロに関する脅威が高まる中，IDS を用いてネットワークシステムを監視する SOC(Security Operation Center)の構築が進められている[1]．SOC では，IDS から出力されるイベントの頻度に注目して，通常と異なる特徴を持ったイベントの検出に努めている．しかしながら IDS の多くは，誤検知，繰り返し検知，多

重検知，対策済み検知など，冗長なイベントを多量に含むため，異常検出には難しい課題がある．個々のイベントを効率的に分析する手法として，検知時刻，IP，Port 情報などを基に類似のイベントをグループ化するクラスタリング分析[2][3]や相関分析[4]などが提案されている．著者らも，長期間のイベント頻度の統計分布を用いて短期間のイベント頻度の異常度を算出して，危険なイベントを順位付けするシステムを提案してきた[5].

昨今, Slammer ワームや Blaster ワーム[6]など，急速にインターネット全域に拡大する攻撃が深刻な問題になっている．SOC 運用において，ネットワークサービスに影響を与える大規模な攻撃状況やワーム感染ホストによる局所的な影響を把握することは重要な課題であり，拡散規模を測定する研究がなされている[7]．これは，注目する攻撃イベントの出力特性に関する情報理論的な曖昧度を情報エントロピー[8]を用いて評価する手法を提案している．例えば，多数の Destination IP が記録される攻撃の場合，その曖昧度を攻撃規模とみなして評価している．しかしこの研究は，個々の攻撃特性を評価したに過ぎず，ネットワークサービスに影響を与える攻撃を検出するには至っていない．

そこで本研究では，IDS イベント全体の情報エントロピー値の変化の程度に注目した分析手法を提案する．これは，長期間の情報エントロピー値の統計分布の平均と標準偏差を用いて，短期間の情報エントロピー値の乖離の程度を評価する手法である．本手法について実運用されている IDS ログを用いて局所的攻撃の検出率に関する評価を行った結果，未検出率ならびに誤検出率を低減できること，従来からのイベント頻度を用いた異常検出手法と組み合わせることで確実に検出できることを示す．

以下 2 章において，既存の IDS ログ分析システムとその課題について説明する．3 章で，IDS イベントから算出される情報エントロピー値の異常性を評価する手法について提案する．4 章で，本手法の評価を行い，有効性を確認する．5 章で，本異常検出の結果を攻撃概況指標へ適用することを提案して，最後に 6 章でまとめる．

2. IDS ログ分析システムとその課題

ここでは，本研究のアルゴリズムを適用する基盤システムとして，著者らが開発してきた広域監視のための IDS ログ分析システムについて説明する[5]．また，昨今の広域監視における課題を取り上げ，その要件を整理する．

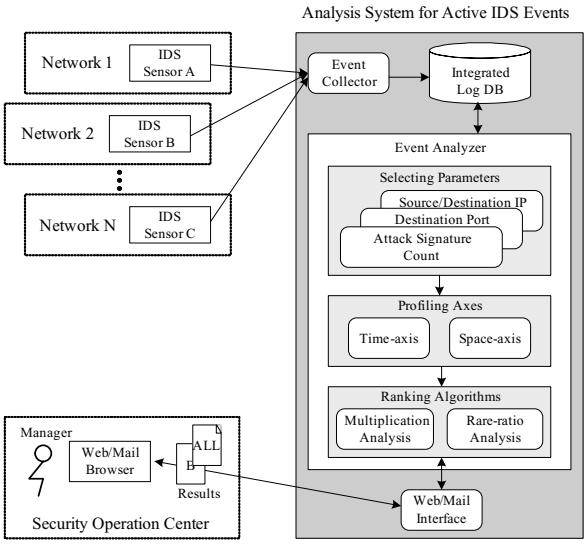

図 1. 広域監視のための IDS ログ分析システム

2.1. IDS ログ分析システム

図 1 に, IDS ログ分析システムの構成を示す．本システムは，各地のネットワークに設置した IDS のログを収集・統合管理して，イベント頻度に注目した統計的な分析処理を行い，分析結果を Web や Mail で監視者に通知する．

イベント分析部で扱うイベントは，各種 IDS に共通して含まれる異常検出に適した以下の 4 つのパラメータである．

- Attack Signature
- Destination Port
- Source IP / Destination IP

これら 4 つのパラメータに関して，時間軸上と空間軸上でのイベント頻度の分布を算出して，通常と異なる頻度の乖離値を評価している．時間軸上での乖離の算出は，長期間で検知されたイベント頻度の平均値に対する短期間で検知されたイベント頻度の倍率で評価する比率分析と，長期間のイベント頻度の平均値と標準偏差を用いて短期間のイベント頻度の統計的な信頼区間の補集合を評価する稀率分析の二通りである．

2.2. SOC 運用の課題と要件

日々新たなワームや攻撃ツールが現れており，分析対象のイベント種別が膨大な数になっている．例えば，クラス C 規模のネットワークを監視している IDS の場合，Attack Signature では数百種類，Destination Port，Source IP，

Destination IP では数千種類のイベントが日々検知されている．このような中，多量の攻撃トラヒックを集中させてネットワークサービスを停止させる攻撃（DDoS: Distributed Denial of Service）や，新規ワームの急速な拡大など，SOC には攻撃の偏りや拡散規模を的確に把握して，迅速な対応を図ることが求められている．

個々のイベント頻度の変動に注目した従来の分析手法は，異常なイベントを的確に抽出できるものの，各種攻撃による総合的な影響を把握するには適していない．また，頻度の小さなイベントの変動が高く評価されてしまいがちで，誤検出が多数発生する問題がある．

以上の背景と問題点より，

[要件 1] 多種多様な攻撃を纏めて評価できること
[要件 2] 攻撃の偏りや拡散規模を把握できること
[要件 3] 分析結果は簡易な指標であること
[要件 4] 確実に検出できること
[要件 5] 誤検出を低く抑えること

が IDS ログ分析の要件として挙げられる．

3. 情報エントロピー値の異常分析

ここでは要件 1,2 を考慮して，多種多様なイベントから攻撃の偏りや拡散規模を総合的に評価する手法として，ログの曖昧度の変化に注目した異常分析手法を提案する．

3.1. 攻撃モデル

まず初めに，本研究で分析対象になる攻撃モデルについて定義する．図 2 に，2.2 節で説明した 4 つのパラメータに注目したときの，A.ワーム/IP スキャン攻撃，B.DDoS 攻撃，C.不正通信を行うアプリケーション攻撃の様子を示す．各攻撃モデルでは，1 対多，多対 1，1 対 1 のイベントが多量に記録される．”1”側のパラメータに注目すると，イベントに偏りが発生して曖昧度が低くなる．”多”側に注目すると，多種類のイベントが記録されて曖昧度が高くなる．

3.2. 情報エントロピーの適用

多種多様な攻撃モデルの影響を総合的に把握するために，各パラメータに含まれる全てのイベントのばらつきを一括して情報エントロピーで評価する手法について提案する．

ある単位時間中に検知されたイベントが n 種類あったとき，イベント i が全イベントに占める割合を pi とする．このときの全イベントの曖昧度を表す情報エントロピー値 H は，(1)式に従い算出される．

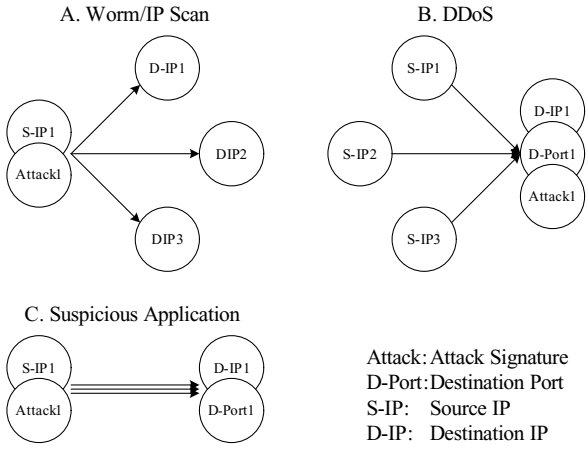

図 2．本研究で対象にする攻撃モデル

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \qquad (1)$$

3.3. 異常分析

(1)式で算出される情報エントロピーの値がどの程度異常であるかを判断するために，時間軸上での情報エントロピー値の統計分布を用いた評価手法を提案する．

評価対象にする短期間を単位時間としたときに，その情報エントロピー値 Hs が，それよりも過去の長期間の単位時間あたりの情報エントロピー値の平均 μ に対して，標準偏差 σ に換算してどの程度かけ離れているかを，その信頼区間の補集合で評価する．この様子を図 3 に示す．短期間の情報エントロピー値が平均よりも小さな場合（左側）の信頼区間の補集合を下側稀率と呼び，平均よりも大きな場合（右側）の信頼区間の補集合を上側稀率と呼ぶことにする．

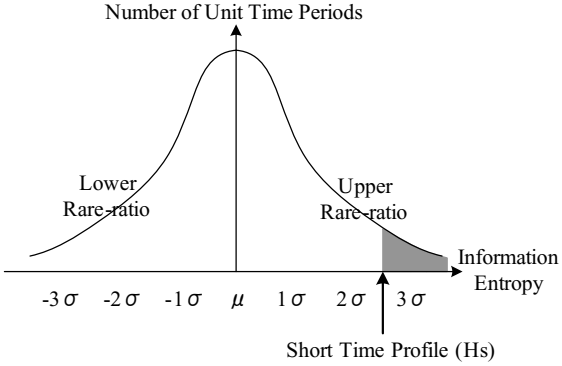

図 3．平均と標準偏差を用いた異常分析

m 個の単位時間からなる長期間の情報エントロピー値の平均は，

$$\mu = \frac{\sum_{k=1}^{m} H_k}{m} \qquad (2)$$

と求められ，その標準偏差は，

$$\sigma = \sqrt{\frac{\sum_{k=1}^{m} (H_k - \mu)^2}{m}} \qquad (3)$$

と求められる．

本研究では，長期間の情報エントロピー値の統計分布は，(4)式の確率密度関数で表される正規分布に従うものと仮定する．

$$f(x) = \frac{1}{\sqrt{2\Pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad (4)$$

これより，下側稀率 *Rl* と上側稀率 *Ru* は，それぞれ式(5)と(6)のように導出される．

$$R_l = \int_{-\infty}^{H_s} f(x)dx \qquad (5)$$

$$R_u = \int_{H_s}^{\infty} f(x)dx \qquad (6)$$

3.4. 実装例

図 4 に，2003 年 8 月の Attack Signature の情報エントロピー値の変動グラフと，グラフ最終日の稀率を算出するシステムの出力例を示している．図では，8 月 18 日の情報エントロピー値が急減している．この日の情報エントロピー値は 0.33 であり，その下側稀率は 0.12%であった．この日は，Blaster ワーム[6]に感染したホストの活動が多量に記録された日であり，局所的攻撃の影響が顕著に現れている．このように、情報エントロピー値の変動に注目することで，攻撃の偏りの程度を容易に把握することができる（要件 2,3 の達成）．

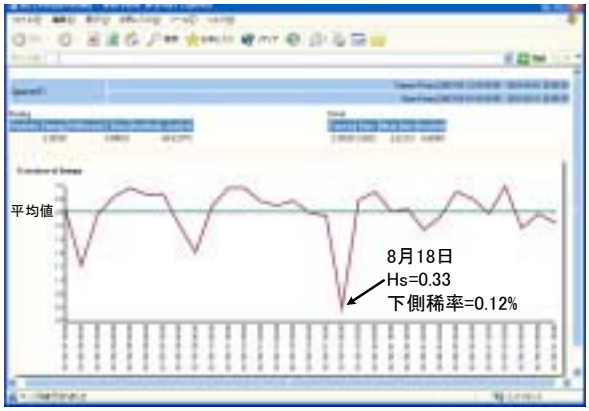

図 4. Attack Signature の情報エントロピー値の変動

4. 評価

ここでは，情報エントロピー値の稀率に注目した監視の妥当性を確認するために，実際に運用されている IDS ログを用いて，局所的攻撃の未検出率と誤検出率に関して評価を行う．

4.1. 評価条件

評価は，クラス C 規模のネットワークを監視している IDS ログを用いる．評価対象の短期間を 1 日とし，長期間を 1 ヶ月とする．IDS ログは，2003 年 7 月から 2003 年 12 月までのものを用い，7 月のログは 8 月の単位時間を評価するためにのみ利用する．稀率を算出する全日数は，8 月 1 日から 12 月 31 日までの計 153 日となる．

予備調査として 153 日間の全ての IDS ログについて，著者らの手動による詳細な解析を行い，図 2 に示したネットワークサービスに大きな影響を及ぼす攻撃を計 13 日確認した．この 13 日の局所的攻撃のうち，7 日がワーム/IP スキャン攻撃であり，残りの 6 日が不正通信を行うアプリケーション攻撃であった．

本研究の情報エントロピー値の稀率を評価するために，従来研究のイベント頻度の稀率[5]と比較する．ここで、従来研究は個々のイベントの異常性を稀率で評価する手法であるが，ここでは全てのイベントを一つに纏めた総頻度を評価する手法へと応用して、その稀率を用いることにする．

4.2. 本方式と従来方式の比較

Attack Signature に注目して評価を行うこととし，局所的攻撃がなされたときに，本方式では下側稀率が小さくなり，従来方式では上側稀率が小さくなることに注目して，検出率の評価を行う．検出のための閾値を 5%に設定し，これよりも稀な日に正しく局所的攻撃が検出されていることを確認する．

表 1 に，局所的攻撃が記録された日，本方式の下側稀率が閾値より小さくなった日，従来方式の上側稀率が閾値より小さくなった日のうち，いずれかに該当する日の局所的攻撃の有無と各方式の検出の様子を示す．表中の”A”は図 2 の A.ワーム/IP スキャン攻撃日を、”C”は C.不正通信を行うアプリケーション攻撃日を、”N”は誤検出日をそれぞれ表す。また表 2 に，正検知，未検出，誤検出の回数と率を纏めた結果を示す．

表 1 および表 2 より，本方式は従来方式に比べて未検出率ならびに誤検出率が小さく，優れた方式であることがわかる．これは、頻度は曜

日や時間によって大きく変動する傾向があるのに対して、情報エントロピー値は各イベントの占める割合に注目することで、曜日や時間の影響を受けにくい安定した傾向を持ち、局所的攻撃によるイベントの偏りを的確に検出できるためである。

表 1. 閾値 5%のときの局所的攻撃の検出特性

	大規模攻撃	本方式（下側稀率）	従来方式（上側稀率）
N-1 日目	無	-	検出
A-1 日目	有	検出	検出
N-2 日目	無	-	検出
N-3 日目	無	検出	-
A-2 日目	有	-	検出
A-3 日目	有	検出	-
A-4 日目	有	検出	-
C-1 日目	有	検出	検出
A-5 日目	有	検出	-
N-4 日目	無	検出	-
N-5 日目	無	検出	-
N-6 日目	無	検出	-
C-2 日目	有	検出	-
C-3 日目	有	-	検出
N-7 日目	無	-	検出
N-8 日目	無	-	検出
A-6 日目	有	検出	-
A-7 日目	有	検出	-
C-4 日目	有	検出	-
C-5 日目	有	検出	-
C-6 日目	有	検出	-
計（回）	有 13 無 8	該当 15	該当 8

表 2. 表 1 の検出回数および検出率のまとめ

	正検出	未検出	誤検出
本方式	11 回 85%	2 回 15%	4 回 27%
従来方式	4 回 31%	9 回 69%	4 回 50%
組合せ方式（OR 条件）	13 回 100%	0 回 0%	8 回 38%

しかし，本方式のみでは未検出になる場合もあるため，本方式と従来方式のどちらか一方でも検出された場合を検出日とみなす組合せ方式についても考える．組合せ方式では，誤検出率は従来方式よりも小さく，未検出率は 0%になり，確実に局所的攻撃を検出できていることがわかる(要件 4,5 の達成)．

さらに表 1 より，本方式と従来方式の両方で検出された日が 2 回あるが，2 回とも局所的攻撃が行われた日に一致していることがわかる．

4.3. パラメータ相互評価

ここでは，のべ 7 日で検知された A.ワーム/IP スキャン攻撃に関して，Attack Signature，Source IP，Destination IP の相互関係に注目する．本方式の結果としては，Attack Signature と Source IP では下側稀率が小さくなり，Destination IP では上側稀率が小さくなることが予想される．表 3 に，該当する日の稀率の結果を示す．表中の”下”は下側稀率を表し，“上”は上側稀率を表す．

表 3. ワーム攻撃における稀率

	Attack	S-IP	D-IP
A-1 日目	下 0.12%	下 0.07%	上 64.25%
A-2 日目	下 50.23%	下 11.64%	上 55.87%
A-3 日目	下 0.74%	下 0.39%	上 18.68%
A-4 日目	下 4.77%	下 5.36%	上 84.15%
A-5 日目	下 0.23%	下 12.60%	上 94.76%
A-6 日目	下 0.41%	下 0.30%	上 0.10%
A-7 日目	下 2.57%	下 0.80%	上 38.98%

表 3 より，Attack Signature と Source IP の稀率は，2 日目を除いてほぼ同じように連動していることがわかる．これは，同じ Source IP から同じ Attack Signature が繰り返し検知されてイベントが偏るためである．Destination IP の上側稀率は 3, 6, 7 日目に小さくなっており攻撃が多数のホストに拡散している様子がわかる．逆に 1, 2, 4, 5 日目については，感染したホストが選んだ伝染先の Destination IP が隣接する IP であったため，日々検知されている Destination IP に重複して攻撃が行われることで，情報エントロピー値が下がっている．6 日目以外は，拡散規模としては特に大きな攻撃ではないことがわかる(要件 2 の達成)．

5. 攻撃概況指標への応用

4 章の評価結果より，IDS イベントに関して本方式の情報エントロピー値の稀率と，従来方式の総頻度の稀率を用いて監視することで，ネットワークサービスに影響を与える攻撃を確実に検出できることがわかった．ここではこれらの稀率を，攻撃概況を表す指標へ適用することを提案する．

5.1. 既存の攻撃概況指標

セキュリティ情報を発信しているサイトの中には，ネットワーク運用者や一般利用者向けに，攻撃概況に関する危険度指数を発表している．例えば，”InfoCon [1]”，”ThreatCon [9] [10]”，”AlertCon [11]”などがある．これらの指標は，セキュリティホール情報，IDS イベントの頻度，実社会の情勢などを考慮して，1 回/日の頻度で導出されている．しかしながら，この導出過程には，SOC 運用者の主観が入りやすく基準が明確になっていない問題と，更新が日単位でしか発表されない遅延の問題がある．

攻撃概況指数には，社会情勢などの曖昧なパラメータを考慮しないこと，普遍的な算出基準で自動的に算出できることなどが求められる．

5.2. 情報エントロピー値の稀率の適用

5.1 節の要件を考慮して，本方式の稀率ならびに従来方式の稀率をインターネットの攻撃概況指標として利用することを提案する．

図 2 の IDS ログ分析システムは，定期的に各地の IDS ログを収集している．このログを収集するタイミングで，Attack Signature に関する本方式の稀率と従来方式の稀率を算出する．

攻撃の危険度を 4 段階に分類することにする．4.2 節の評価では，従来方式よりも本方式が適切にネットワークサービスに影響を与える攻撃を検出できることが示されている．そこで，本方式の稀率を重視するように，表 4 に示す閾値に従い 4 段階の危険度に対応付けることにする．“－”印は稀率 5%以上を表す．最終的に，表 4 の結果を棒グラフで通知する．この棒グラフは，ネットワークサービスに影響を与える攻撃を明確な基準で評価した客観的な指標であり，SOC 運用者にとって理解しやすく普遍的な数値として期待される(要件 3 の達成)．

表 4. 攻撃概況指標の基準値例

レベル	本方式の稀率	従来方式の稀率	危険度
1	－	－	低
2	－	上側 5%未満	中
3	下側 5%未満	－	高
4	下側 5%未満	上側 5%未満	最高

6. おわりに

本研究では，各種 IDS イベントの曖昧度をパラメータ単位で纏めて情報エントロピーで算出し，その異常性を評価する手法を提案した．異常性は，短期間の情報エントロピー値を長期間の統計分布の平均と標準偏差を用いて評価する手法であり，信頼区間の補集合として出力される．これにより，総合的な攻撃の偏りや拡散規模を容易に把握できるようになった．実運用されている IDS ログを用いて評価を行った結果，本手法と従来手法を組み合わせることで，局所的攻撃について誤検出率を小さく抑えながら確実に検出できることを確認した．各地に設置した IDS ログを用いて広域監視を行うことで、大規模攻撃の検出も可能になる。また，これらの稀率を攻撃概況指標へ適用することを提案した．この指標は，普遍的な算出手順であるため，自動的に攻撃概況を評価することができ，SOC 運用者の迅速な対応と情報交換に寄与する．

文　献

[1] Internet Storm Center, http://isc.incident.org/

[2] K. Julisch, "Mining Alarm Clusters to Improve Alarm Handling Efficiency", In Proceedings of the 17th Annual Computer Security Applications Conference (ACSAC2001), December 2001.

[3] F. Cupperns, "Managing alerts in a multi-intrusion detection environment", In Proceedings of the 17th Annual Computer Security Applications Conference (ACSAC2001), pp.22-31, December 2001.

[4] A. Valdes and K. Skinner, "Probabilistic Alert Correlation", The 4th International Symposium on Recent Advances in Intrusion Detection (RAID 2001), pp.54-68, October 2001.

[5] 竹森敬祐，三宅優，中尾康二，菅谷史昭，笹瀬巌，“セキュリティデバイスログ分析支援システムの広域監視への適用”，情処，コンピュータセキュリティシンポジウム 2003（CSSS2003），pp.397-402，2003 年 10 月．

[6] CERT/CC, http://www.cert.org/

[7] 小泉芳，小池英樹，高田哲司，安村通晃，石井威望，“情報エントロピーを用いたネットワーク侵入検知システムログ解析手法の提案”，情処，コンピュータセキュリティシンポジウム 2003（CSS2003），pp.653-658，2003 年 10 月．

[8] C. E. Shannon: “A mathematical theory of communication”, Bell system tech. J., Vol.27, pp.379-423, pp.623-656, 1948.

[9] Deep Sight, http://tms.symantec.com/

[10] Security Focus, http://www.securityfocus.com/

[11] Internet Security Systems, http://www.iss.net/