LLL基底簡約アルゴリズムを利用したDDoS攻撃検知の考案

森田 拓哉[1,a)]　鈴木 彦文[2,b)]　岡崎 裕之[1,c)]

概要：近年，IoT デバイス等の急速な普及により，パソコンやスマートフォン等の従来の接続端末に加えて家電，自動車や工場など，様々な「もの」がインターネットに接続されるようになっている．それらのデバイスの普及に伴い，世界中で様々なサービスが提供され生活を支える重要な基盤となっている．その中でも公開性の高いサービスはサイバー攻撃の標的になりやすく，防衛手段として UTM 機器などの防衛装置を利用されているが，近年これらの防衛装置をすり抜ける攻撃が増加している．その攻撃の中でも特に DDoS 攻撃はサービスそのものを停止させるため社会的なインパクトが大きく，正規の通信に偽装しているため検知することが困難である．そこで本研究では，各通信を多次元空間に存在する点とみなし，点の並びの規則性から DDoS 攻撃の通信を検知することを試みた．点の並びの規則性を調べるために LLL 基底簡約アルゴリズムを用いて通信ログから基底を算出し，その基底のノルムや次元を調べる．MWS データセット [1] と信州大学ネットワークの通信ログの解析結果の比較を行い，基底の特徴から攻撃による通信を検知する手法の有効性を検証した．

キーワード：基底，ログ解析，MWS，UTM，DDoS 攻撃

Devising a DDoS Attack Detection Method Using LLL Basis Reduction Algorithm

Takuya Morita[1,a)]　Hikofumi Suzuki[2,b)]　Hiroyuki Okazaki[1,c)]

Abstract: DDoS attacks are difficult to detect because they are disguised as legitimate communications. Furthermore, it has a significant social impact because it shuts down the service itself. In this study, we consider each communication as a point in a multidimensional space and try to detect the communication of a DDoS attack from the regularity of the sequence of the points. In order to investigate the regularity of the order of points, the basis is calculated from the communication log using LLL basis simplification algorithm, and then the norm and dimension of the basis are investigated. We compare the results of the analysis of the MWS dataset and the communication logs of Shinshu University network to verify the effectiveness of the method for detecting communication due to attacks based on the characteristics of the base.

Keywords: basis,　log analysis，MWS，UTM，DDoS Attack

1. はじめに

近年，情報通信機器の普及とインターネットの利用率の増加に伴い，様々なサービスの実現が可能となり生活の一部となっている一方で，サービスの重要度に伴いセキュリティに対する意識も高まっている．セキュリティ装置である UTM（Unified Threat Management）機器等が発達す

[1] 信州大学大学院総合理工学研究科工学専攻
Science and Technology Department of Engineering Electrical and Computer, Engineering Division, Shinshu University
[2] 信州大学総合情報センター
Integrated Intelligence Center, Shinshu University
[a)] 19w2127a@shinshu-u.ac.jp
[b)] h-suzuki@shinshu-u.ac.jp
[c)] okazaki@cs.shinshu-u.ac.jp

る一方で，サイバー攻撃の脅威は増加し，攻撃方法も巧妙化している．

サイバー攻撃の中で最も多い手口の一つである DDoS（Denial of Service/Distributed DoS）攻撃 [2], [3], [4] は，既存の UTM 機器において通常のアクセスのように観測されることが多く検知することが難しい．IP アドレスも偽装されたものがほとんどであり，攻撃時間の間隔等も関係してくるので，検知する手段としてトラフィックログの各フィールドを一つずつ見ても判断することができない．通常のトラフィックと攻撃した際のトラフィックログを判別するには複数のトラフィックログ，複数のフィールドの何かしらの関係性を発見する必要がある．そこで，本研究では DDoS 攻撃を含んでいる部分集合を張る格子基底[*1] を見つけることで DDoS 攻撃を特定することを試みた．

具体的には図 1 にあるように点の集合を格子と見立ててトラフィックデータを扱う．この点の集合から DDoS 攻撃である部分集合を見つけるということをする．図 1 の集合は図 2 のような基底を用いて線形変換することで表現することができる．この基底が変化すると表現する空間も変化する．図 3 のような緑の点の集合（部分集合）のみを表現する基底は図 3 中のような基底となる．このように特定の部分空間を表現する格子基底を求めることによって DDoS 攻撃を特定していく．今回はネットワークのフローを表現する 5-tupple[*2] という 5 つの要素を用いるのでイメージとしては図 4 のようなものとなる．5-tupple のそれぞれ 5 つの要素が基底となり 5 次元空間全体を表現している．この 5 次元空間の中から攻撃トラフィックを含む部分空間のみを表現する基底を攻撃トラフィックの LLL 簡約基底で近似できるかどうかの実験を行った．

本研究では実際の環境に近いトラフィックを取得するため，実際に信州大学のサーバに対して外部から SINET 経由で DDoS 攻撃を実施し，UTM 機器からトラフィックログを取得し解析を行った．信州大学の UTM 機器のトラフィックログの解析結果と MWS データセットの DRDoS 攻撃データの解析結果の比較を行い，基底の特徴から攻撃による通信を検知する手法の有効性を検証した．

2. 研究の概要

本研究は DDoS 攻撃という既存のセキュリティ装置では検知することが困難であるサイバー攻撃を検知するため

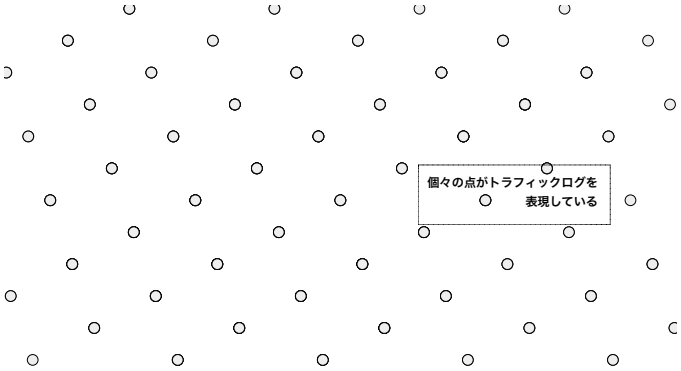

図 1 全てのトラフィックを表す集合

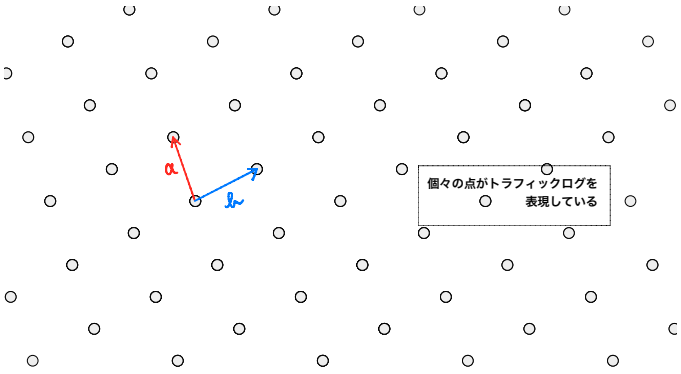

図 2 全てのトラフィックを表現する基底

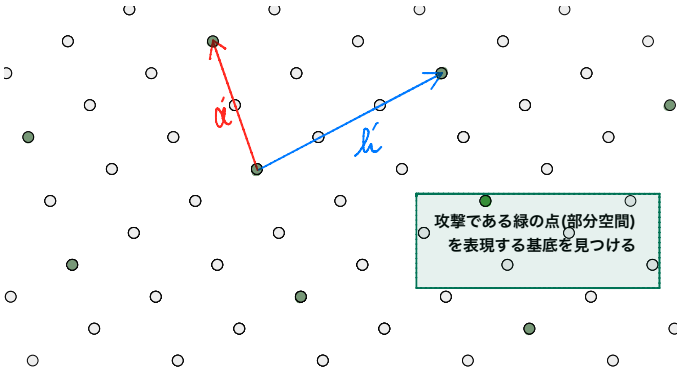

図 3 特定の部分空間を表現する基底の例

に，信州大学のサーバーに攻撃を行った．攻撃後，UTM 機器からトラフィックログを取得し，以下の二つのデータに分ける．

- DDoS 攻撃時のトラフィックだけを抽出した
  トラフィックログ
- DDoS 攻撃時のトラフィックを取り除いた
  非攻撃時のトラフィックログ

上記のそれぞれのデータから基底を計算し，その特徴を見ることで攻撃を検知できないか検証する．

基底を求めるために本稿では SVP（最短ベクトル問題）の近似解を求めることができる LLL 基底簡約アルゴリズムを用いた.LLL 基底簡約アルゴリズムを用いて算出した

[*1] ここでいう基底とはベクトルの集合 $\{\vec{v}_1,\cdots,\vec{v}_n\}$ において，以下の 2 つの条件を満たすものを指す．
条件 1．それらの一次結合で全てのベクトルを表現できる
条件 2．それらは一次独立である

[*2] ネットワークのフローを表現する以下の 5 つの要素
- 送信元 IP
- 宛先 IP
- 送信元ポート
- 宛先ポート
- プロトコル番号

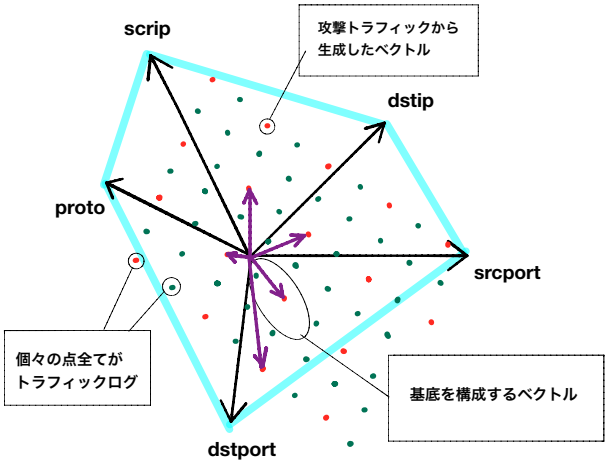

図 4 5-tupple をベクトルに変換したものと基底

ベクトルを LLL 簡約基底といい本稿ではこれを基底として取り扱う.

基底を計算する際に利用するログのフィールドは 5-tupple と呼ばれるネットワークのフローを表現する送信元 IP，宛先 IP，送信元ポート，宛先ポート，プロトコル番号の 5 つとする.

トラフィックを集合として考えると DDoS 攻撃のトラフィックは正常なトラフィックの部分集合と考えられるので算出された基底に表れる特徴として以下のようなものがあげられる.

- DDoS 攻撃時のトラフィックログの基底のノルムが正常なトラフィックのものと比較して長くなっている
- DDoS 攻撃時のトラフィックログの基底を構成するベクトルの次元が正常なトラフィックのものと比較して低くなっている

DDoS 攻撃時のトラフィック群から生成された基底に上記の二つの特徴が見られれば以下の二つの基底の特徴を比較する.

- DDoS 攻撃時のトラフィックだけを抽出したトラフィックログ
- MWS データセットの DRDoS 攻撃のログ

取得経路の違う二つのデータの解析結果を比較して基底を用いて DDoS 攻撃を検知する手法の有効性を調べる.

2.1 UTM トラフィックログの取得環境

本研究にて利用した UTM トラフィックログの取得環境は 攻撃サーバ 1 台・被攻撃サーバ 1 台および UTM 機器 1 台から構成される. 本環境の概略図を図 5 に示す. 攻撃サーバから被攻撃サーバに対して攻撃ツールをもちいたパケット送出を行う. この二者間に UTM 機器を配置し，ネットワーク上を流れるセッション一つ一つをログとして残すものとなっている. 攻撃パケットの生成にはオープンソースのツールを用いることとした. 生成したい攻撃の種類としては http flood，tcp syn flood，udp flood および DNS Amplification の 4 種類があげられる.tcp syn flood および udp flood の実施には loic[*3]，t50[*4] と hping3[*5] を，http flood の実施には loic を，DNS Amplification の実施には saddam[*6] をそれぞれ用いることとした. 今回取り扱うトラフィックは表 1 に示したように信州大学サーバーに対して saddom を用いて 3 つのアドレスから 5,133 件の DNS Amplification 攻撃を行ったものを利用する.

表 1 今回取り扱う信州大学への DDoS 攻撃の概要
Table 1 Overview of the DDoS attack on Shinshu University

ツール	実施時間	手法	ホスト数	件数
saddam	15:03:01~15:03:33	DNS Ampflication	3	5,133

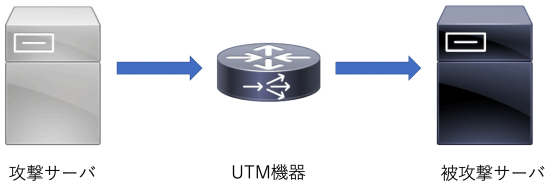

図 5 UTM ログデータ取得環境

2.2 MWS の DRDoS 攻撃のデータセット概要

今回用いる MWS のデータセットの DRDoS 攻撃データは図 6 に示すように 192.168.11.5 に対して多数の送信元から攻撃を行ったものである. 提供されているものは pcap ファイルなので今回は Python の dpkt[*7] というライブラリを用いて必要な要素を抽出した.

3. 基底の特徴を利用した解析方法

本研究では DDoS 攻撃のトラフィックの規則性を基底を用いて明らかにする.

図 4 のようにトラフィックログの 5-tupple をベクトルに変換したものから基底を計算する. 算出された基底の長さ，次元を調べ，DDoS 攻撃のトラフィックログがどのような規則性を持って存在しているのかを調べていく.

今回は 5-tupple から生成したベクトルがトラフィックログが存在しうる多次元空間のどの場所にどのような規則性をもって存在しているかを調べたいので，5-tupple の全ての値をそれぞれのフィールドがとりうる最大値で割ることによって正規化したものを利用した. 正規化したベクトルから LLL 基底簡約アルゴリズムを用いて計算したものを基底とみなし，その特徴を見ていく.

*3 github.com/neweracracker/loic
*4 github.com/foreni- packages/t50
*5 linux.die.net/man/8/hping3
*6 github.com/offensivepython/saddam
*7 pcap 解析用のライブラリ

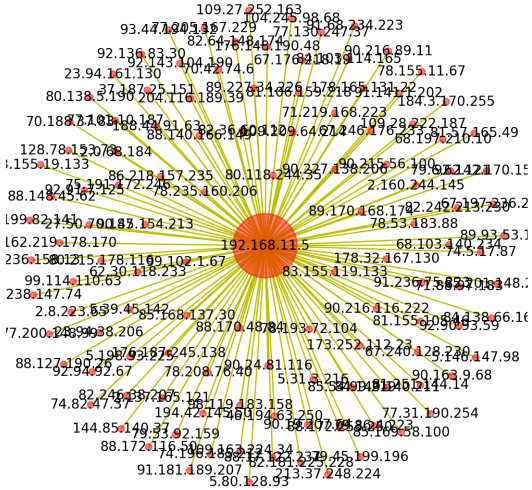

図 6 MWS の DRDoS 攻撃のデータセットのグラフ

3.1 LLL 基底簡約アルゴリズムの利用

格子 L の基底 $\{b_1,\cdots,b_n\}$ が与えられたとき，それぞれの基底ベクトル が短くかつ互いの基底ベクトルが直交に近い同じ格子 L の基底に変換する操作を格子基底簡約（lattice basis reduction）と呼ぶ．LLL（Lenstra-Lenstra-Lovász）基底簡約は最も有名かつ代表的な格子基底アルゴリズムで，n 次元格子上の SVP を近似的に解く効率的なアルゴリズムである [4]．LLL 基底簡約アルゴリズムは一般次元の格子状の最短ベクトルを必ず見つけるわけではなく．具体的には，LLL 基底簡約アルゴリズムは n 次元格子 L の第 1 次逐次最小 λ_1（L）にある指数因子をかけた値より小さいノルムを持つ格子ベクトルを効率的に見つけ出すアルゴリズムである．つまり，LLL 基底簡約はある指定因子による近似版の最短ベクトル問題を効率的に解くアルゴリズムである．

本稿では DNS Amplifcation のトラフィックログと正常と思われるトラフィックログそれぞれに対して sagemath[*8]の機能の一つである PARI/GP に含まれる LLL 基底簡約アルゴリズムを用いて LLL 簡約基底を生成する機能を使用した．生成した LLL 簡約基底を基底ベクトル（格子ベクトル）とみなして長さ，次元を調べる．

4. 解析結果の比較，評価

最初に信州大学への DDoS 攻撃した際のトラフィックから計算した基底を表 2 に示す．表 2 を見ると基底は 5 次元ベクトルであり，各ベクトルが 2 次元で構成されている．ノルムも 5 本中 3 本が顕著に大きくなっていることがわかる．

[*8] http://www.sagemath.org/

表 2 5-tupple から生成した基底 (信州大学への DDoS 攻撃)
Table 2 Basis generated from 5-tupple (DDoS attack on Shinshu University)

	Basis
5-tupple	0, 1, -2709228, 26454016;
	30915886641053, 18384410234110764500, 256204388056399079868587749 1252, 24489930457462760162108849;
	0, 0, 0, 1;
	-24039909392479820, -1429554847944507411 5844, -19922217810997226187 78638774368448, -19043144906727610906984140804;
	0, 0, 1, 0;

かる．

続いて，5-tupple の要素から組み合わせを変えて基底を生成していく．表 3 は得られた基底を構成するベクトルが 2 次元ベクトルだったものをまとめた表である．表 4 は得られた基底を構成するベクトルが 1 次元ベクトルだったものをまとめた表である．この結果から基底を構成するベクトルのノルムが一部顕著に長くなっていることがわかる．また，基底を構成するベクトルの次元において最大で 4 次元，最低で 1 次元になっている．

続いて，信州大学の非攻撃トラフィックから生成した基底を見ていく．出力として得られた結果を表 5 に示す．表 5 を見ていくと基底を構成するベクトルのノルムが最大で 11 と小さいことがわかる．さらに，基底の次元と基底を構成しているベクトルの次元が一致している．．これらの結果より，以下の二つのような仮説を立てることができる．

- 攻撃時のトラフィックは基底のノルムが極端に長くなる．
- 非攻撃時のトラフィックは基底の次元と基底を構成するベクトルの次元が一致している．

続いて先述した二つのデータの規則性が他のデータにも当てはまるを確認するために MWS データセットの DRDoS 攻撃のログを基底を用いて解析する．MWS データセットの DRDoS 攻撃のログから 5-tupple を用いて基底を計算したものを表 6 に示す．

MWS データセットの DRDoS 攻撃のログも信州大学の DDoS 攻撃のトラフィックと同様に基底のノルムが顕著に長くなっていることがわかる．さらに，基底の次元と比較すると基底を構成している次元は”dstip 無し & proto 無し”の場合を除いて低くなっている．これらの比較結果から，DDoS 攻撃を基底のノルムの長さで判断することが有効であると思われる．さらに，基底の次元と比較して基底を構成している次元が小さい場合は攻撃であると考えることができる．

表 3 各ベクトルが 3 次元以上で構成されている基底を生成する組み合わせ (信州大学への DDoS 攻撃)

Table 3 Combinations that produce a basis where each vector consists of more than three dimensions (DDoS attack on Shinshu University)

	Basis
dstip 無し	1129, -1765, -2709768, 26454258;
	0, 0, 0, 1;
	-1106720, 1730169, 2656292330, -25932202941;
	0, 0, 1, 0;
dstport 無し	-45821, 10827, -2735047, 26430070;
	57764, -13649, 3447922, -33318928;
	0, 0, 0, 1;
	0, 0, 1, 0;
srcip 無し	30915886641053, 256209368811394225032742057294O, -27825611290217056166866950O2;
	0, 0, 1;
	-24039909392479820, -199226051099292595106847156076360O, 216369396735631785715936443110O;
	0, 1, 0;
srcport 無し	0, 1, -2709228;
	30915886641053, 18384410234110764500, 256204388056399079868587749125 2;
	-24039909392479820, -142955484794450741158 44, -199222178109972261877863877436844 8;
	0, 0, 1;
proto 無し	0, 1, 26454016;
	30915886641053, 18384410234110764500, 244899304574627601621088 49;
	0, 0, 1;
	-24039909392479820, -142955484794450741158 44, -190431449067276109069841408 04;
dstip 無し & proto 無し	1129, -1765, 26454258;
	0, 0, 1;
	-1106720, 1730169, -25932202941;

表 4 各ベクトルが 2 次元以下で構成されている基底を生成する組み合わせ (信州大学への DDoS 攻撃)

Table 4 Combinations that produce a basis where each vector consists of less than or equal to two dimensions (DDoS Attacks on Shinshu University)

	Basis
srcport 無し & proto 無し	0, 1;
	30915886641053, 18384410234110764500;
	-24039909392479820, -14295548479445074115844;
srcport 無し & srcip 無し	30915886641053, 256209368811394225032742057294O;
	-24039909392479820, -199226051099292595106847156076360O;
	0, 1;
dstport dstip	30915886641053;
	-24039909392479820

5. 今後の課題，まとめ

本研究では DDoS 攻撃のトラフィックの規則性として，二種類のトラフィックデータにおいてノルムが長く，次元が低い基底が生成された．このことから基底を用いて攻撃検知をすることが有効であることがいえる．また今回は DDoS 攻撃に対してのみを対象にして検証を行ったが，ブルートフォースや辞書攻撃などの多くの通信を行って認証を突破する攻撃や，認証に対しての負荷攻撃などの検知も行えるのではないかと考えられる．今後はそのような攻撃を実際に行い，基底を用いた検知の汎用性を検証する必要がある．また，判別するための基準となる基底のノルムを調べ，何件のログから基底を生成するかを調べる必要がある．

謝辞 本研究の一部は，日本電信電話株式会社 ネットワークサービスシステム研究所と信州大学における共同研究「ネットワークセキュリティにおける機械学習適用の研究」の支援を受けて実施した．また，本研究で利用したネットワーク装置において，信州大学ネットワークの管理・運用に当たっている，東日本電信電話株式会社様のご協力を得て本研究を実施いたしました．ここに感謝の意を表します．

参考文献

[1] 神薗 雅紀, 秋山 満昭, 笠間 貴弘, 村上 純一, 畑田 充弘, 寺田 真敏, ”マルウェア対策のための研究用データセット〜MWS Datasets 2015〜”，情報処理学会，2015-CSEC-70, No.6，pp.1-8，2015-6-25.

[2] 寺田真敏，”DoS 攻撃:1. DoS/DDoS 攻撃とは”，情報処理学会，情報処理 54(5)，428-435，2013-04-15

[3] 高倉弘喜，”DoS 攻撃:2.1 DoS/DDoS 攻撃観察日記 (1)〜DDoS は身内からもやってくる〜”，情報処理学会，情報処理 54(5)，436-444，2013-04-15

[4] 齋藤衛，”DoS 攻撃:3.1 DoS/DDoS 攻撃対策 (1)〜 ISP における DDoS 対策の現状と課題〜”，情報処理学会，情報処理 54(5)，468-474，2013-04-15

[5] 青野良範，安田雅哉，格子暗号解読のための数学的基礎，近代科学社，東京，2019/9/30

[6] 基底ベクトル・双対基底ベクトルと反変成分・共変成分（計量テンソル・クリストッフェル記号・共変微分とは何か,http://fnorio.com/0180covariant_contravariant/covariant_contravariant.html#2, 参照 Oct.24,2019.

[7] 森田拓哉，鈴木彦文，岡崎裕之，”LLL アルゴリズムを用いた UTM トラフィックログの解析”，電子情報通信学会，信学技報，vol. 119，no. 297，NS2019-124，pp. 27-30，2019 年 11 月.

[8] 有元 康一，平野 康之，”虚二次体における LLL 格子基底簡約アルゴリズム (高度情報化社会に向けた数理最適化の新潮流)”，数理解析研究所講究録 (2108):2019.4 p.115-123

表 5 信州大学の非攻撃トラフィックから生成した基底

Table 5 The basis generated from non-attack traffic at Shinshu University

	Basis
5-tupple	0, 0, -1, 1, 5;
	-1, 0, 1, 0, 2;
	1, 0, 0, 0, -11;
	0, 1, 0, -1, 8;
	0, 0, 0, 0, 1;
proto 無し	0, 0, -1, 1;
	-1, 0, 1, 0;
	1, 0, 0, 0;
	0, 1, 0, -1;
dstip 無し	0, 1, -1, 6;
	0, 0, 1, -10;
	1, -1, -1, 8;
	0, 0, 0, 1;
srcip 無し	-1, 0, 1, 5;
	1, 0, 0, -11;
	0, 1, -1, 9;
	0, 0, 0, 1;
srcport 無し	0, -1, 1, 4;
	0, 1, 0, -10;
	1, 0, -1, 7;
	0, 0, 0, 1;
dstport 無し	0, -1, 1, 5;
	-1, 1, 0, 4;
	1, 0, 0, -11;
	0, 0, 0, 1;
srcport 無し & proto 無し	0, -1, 1;
	0, 1, 0;
	1, 0, -1
srcport 無し & srcip 無し	0, 1, -7;
	1, -1, 7;
	0, 0, 1;
dstport dstip	0, 1;
	1, -1;
dstip 無し & proto 無し	0, 1, -1;
	0, 0, 1;
	1, -1, -1

表 6 5-tupple から生成した基底 (MWS)

Table 6 The basis generated from the 5-tupple (MWS)

	Basis
5-tupple	0, 1, 0;
	977867617436294580603633661155723, 4835879451923920456644804801608, 979522222607857084885626560049750;
	0, 0, 1;
	-25383406839758777819892457012059115 62, -125529359360559327223354956465779 5, -254263569441287858303319430942387 6879;
	1, 0, 1;

表 7 各ベクトルが 3 次元で構成されている基底を生成する組み合わせ (MWS)

Table 7 Combinations that produce a basis where each vector is composed of three dimensions (MWS)

	Basis
proto 無し	0, 1, 0;
	-795154646420741, 4835879451923920456644804801608, 16546051715625042819928988940 27;
	0, 0, 1;
	2064055863056240854, -125529359360559327223354956465779 5, -429501043700080104394860821796 5317;
dstip 無し	0, 1, 0;
	0, 0, 1;
	801078014148013617, -974380700771655715243446366811235, -333386173620616299926425417099 4794;
	-3483691925539915, 423734283019900306659027935 7040, 144981475039483094379563169879 25;
dstport 無し	0, 1, 0;
	-84589255483764295, 192970731003883284269258363220705 55, 145691021677775431801105696431 48495;
	0, 0, 1;
	954882596415442889, -217834276468529747328306298092428986, -164462756243072108264307499425422531;
dstip 無し & proto 無し	0, 1, 0;
	0, 0, 1;
	1, -1043, -379;

表 8 各ベクトルが 2 次元で構成されている基底を生成する組み合わせ (MWS)

Table 8 Combinations that produce a basis where each vector is composed of two dimensions (MWS)

	Basis
srcip 無し	977867617436294580603633661155723, 979522222607857084885626560049750;
	0, 1;
	-25383406839758777819892457012059115 62, -254263569441287858303319430942387 6879;
	1, 1;
srcport 無し	0, 1;
	977867617436294580603633661155723, 4835879451923920456644804801608;
	-25383406839758777819892457012059115 62, -125529359360559327223354956465779 5;
	1, 0;
srcport 無し & proto 無し	0, 1;
	-795154646420741, 4835879451923920456644804801608;
	2064055863056240854, -125529359360559327223354956465779 5;

表 9 各ベクトルが 1 次元で構成されている基底を生成する組み合わせ (MWS)

Table 9 Combinations that produce a basis in which each vector is composed of one dimension (MWS)

	Basis
srcip 無し	977867617436294580603633661155723;
&	-25383406839758777819892457012059115 62;
srcport 無し	1;
dstip	-795154646420741;
dstport	2064055863056240854;