全学無線 LAN 利用状況の可視化

Visualization of Wireless LAN Usage Log in University

八切 有市 † 青木 茂樹 † 宮本 貴朗 †
Yuichi Yagiri Shigeki Aoki Takao Miyamoto

1. はじめに

近年，ノート PC やスマートフォン等のモバイル端末の普及に伴い，無線 LAN の利用が増加している．従来，大学等の組織においては教員個人や研究室単位で独自に無線アクセスポイント（以下，AP）を設置し運営していることが多かったが，無線チャネルの枯渇や電波の輻輳の管理の難しさから，今日では大学全体で統括された無線 LAN 環境（全学無線 LAN システム）を整備する動きが盛んになっている．大阪府立大学（以下, 本学）では 2007 年より全学無線 LAN システムを運用しており，無線 LAN システムの管理者（以下，管理者）は，ユーザから無線 LAN の不具合に関する問い合わせがあった場合に，ログを確認することにより全学無線 LAN の状況を把握している．ユーザが円滑に無線を利用できるようにするためには，ユーザからの申告の前に異常を発見したり，異常の兆候を捉える必要がある．そのためには，全学無線 LAN において何らかの異常やその兆候が発生した場合に，管理者に素早くアラートを出し，無線 LAN の状況を分かりやすく可視化して提示できるシステムが望まれている．

無線 LAN の利用状況の分析に関する代表的な研究として，文献 [1,2] が挙げられる．これらの手法では，全学無線 LAN システムから利用ログを収集し，AP ごとの利用方法が変化しているかを分析している．また，ネットワークトラフィックの可視化に関する代表的なシステムとして，独立行政法人情報通信研究機構 (NICT) が開発した NICTER[3] と NIRVANA 改 [4] が挙げられる．NICTER は，広域ネットワークにおけるセキュリティインシデントの迅速な状況把握と原因追求を目的とし，ダークネットに届くパケットを三次元空間または世界地図上に可視化している．NIRVANA 改は，組織内ネットワークを表現した画面上にライブネットのトラフィックを可視化することに加え，各種セキュリティ機器からのアラート集約を行うサイバー攻撃統合分析プラットフォームである．

本論文では，全学無線 LAN 利用ログを単位時間ごとに収集し特徴量を抽出した上で，特徴量の外れ値を検出することで異常を検出する手法を提案する．また，特定の AP の特徴量，その前後の時刻における AP の特徴量，その周囲の AP の特徴量などを容易に把握できる可視化システムを構築することで，管理者が無線 LAN の異常状態を容易に把握できるように支援する．

2. 関連研究

神戸大学における全学無線 LAN 利用ログ情報の解析を行った鳩野の研究 [1] では，神戸大学における無線 LAN システムの利用状況や Learning Commons の利用状況を明らかにすることを目的とし，全学無線 LAN 利用ログを解析する手法を提案している．数年にわたる OS 種別・所属別（教職員，学生等）・接続デバイス数の推移や各学年の在学状況および Learning Commons の利用状況をグラフとして示している．宮崎大学における無線ネットワーク接続のセッションログの解析を行った柳田らの研究 [2] では，宮崎大学における無線 LAN システムの利用状況を明らかにすることを目的とし，無線 LAN のセッションログを解析する手法を提案している．1 年間における所属別・学部別・研究科別の利用割合や 3 年間における学生・職員の無線 LAN 接続率の推移および接続時間の分布を明らかにしている．これらの手法 [1,2] では，ログの短期的な分析には主眼が置かれていないため，無線 LAN の現在の利用状況の把握や異常検知等に応用することは難しい．

一方，ネットワーク全体の現在の状況を可視化する代表的なシステムである NICTER[3] は，インターネット上で時々刻々と変化しているセキュリティインシデントへの迅速な対応を目指したサイバー攻撃観測・分析システムである．NIRVANA 改 [4] は，組織内ネットワークを流れる通信のリアルタイムでの通信・観測や，各種セキュリティ機器からのアラート集約を実現するサイバー攻撃統合分析プラットフォームである．本研究で対象としている無線 LAN 利用ログデータは，NICTER や NIRVANA 改で観測対象としているトラフィックデータとは，得られる情報が異なるため，これらの可視化システムを無線 LAN の利用状況の把握や異常の検知を支援するシステムにそのまま応用することはできない．

本研究では，無線 LAN 利用ログから特徴量を抽出し異常の検出を行うとともに，それらの特徴量を三次元的に描画したアニメーションで表し，わかりやすく提示する手法を提案する．

† 大阪府立大学　Osaka Prefecture University

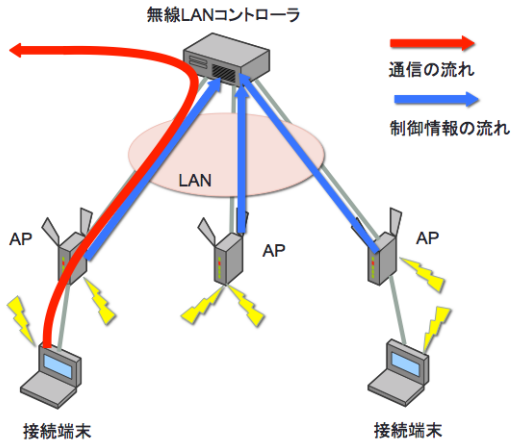

図 1: 無線 LAN システムの構成

3. 提案手法

本研究では，無線 LAN 利用ログから抽出した特徴量から外れ値を検出し，外れ値を検出した場合に AP の状況を三次元的に可視化する手法を提案する．対象とする無線 LAN システムの構成を図 1 に示す．無線 LAN コントローラは，AP の管理，負荷分散の処理，およびハンドオーバーの制御等のために導入しており，無線 LAN コントローラには，各 AP から収集した制御情報等が無線 LAN 利用ログ情報として保存されている．

3.1 無線利用ログ情報の収集方法

無線 LAN コントローラから SNMP を用いて単位時間ごとの無線 LAN 利用ログ情報を取得する．本学では無線 LAN コントローラとして ArubaNetworks 社 (現在はヒューレッド・パッカード株式会社) の無線 LAN コントローラおよび AP[5] を運用している．ArubaNetworks 社の無線 LAN コントローラは，パケットに含まれているフィンガープリンティング情報などを解析することができる．得られた無線 LAN 利用ログには 43 項目の情報が含まれ [6]，その中で異常の発生に伴って変化すると考えられる以下の 7 項目の情報に着目する．

- AP 名
- 接続端末の MAC アドレス，IP アドレス
- 累積接続時間，累積受信バイト数，累積送信バイト数
- SSID
  本学では，ユーザの属性（教職員，学生，外部訪問者など）や利用方法ごとに SSID を分離する運用としており，それらの同定に用いる．

無線 LAN 利用ログには，接続端末ごとに接続開始時からのトラフィック量の累積値が保存されている．本手法ではトラフィック量の変化に注目するために，単位時間でのトラフィック量を算出して用いている．累積接続時間が単位時間以上の接続端末は現時刻での累積送受信バイト数の合計と一時刻前の累積送受信バイト数の合計との差分とし，累積接続時間が単位時間未満であるものは現時刻での累積送受信バイト数の合計とする．

3.2 特徴量抽出

3.1 節で生成した特徴情報およびログ取得時刻をデータベースに登録し，以下に示す特徴量を単位時間ごとに算出する．これらの特徴量はデータベースに登録した特徴情報を基に算出している．ここでは，単位時間ごとに区切った時間を区間と呼ぶ．

- 接続端末種類数
  AP に接続している端末の種類数．MAC アドレス，IP アドレスの情報から算出．
- 総トラフィック量
  AP に接続している全端末のトラフィック量の合計．接続している端末ごとのトラフィック量を合計することにより算出．
- 平均トラフィック量
  AP に接続している端末のトラフィック量の平均．総トラフィック量を接続端末種類数で割ることにより算出．
- SSID 別接続端末数
  AP に接続している，SSID ごとの端末種類数．SSID ごとに MAC アドレス，IP アドレスの情報を調べて算出．求めた値の合計は接続端末種類数と等しい．

3.3 異常の検出

3.2 節で述べたログファイルの最新時刻における特徴量を AP ごとにまとめ，以下の 8 つの判断基準を用いて異常を検出する．

- 同じ曜日の同時刻における接続端末種類数
  同じ曜日の同時刻における接続端末種類数は正規分布に従うと仮定する．外れ値の識別を行う日時の接続端末種類数を $C_0^{<1>}$，その日から n 週間前の同時刻の接続端末種類数を $C_n(n = 1, 2, ..., N_w)$ とする．ここで N_w は注目する週の最大値である．該当時刻における平均 m と不偏分散 V は次式で求められる．

$$m = \frac{\sum_{i=1}^{N_w} C_i^{<1>}}{N_w} \tag{1}$$

$$V = \frac{\sum_{i=1}^{N_w} (C_i^{<1>} - m)^2}{N_w - 1} \tag{2}$$

式 (3) により，外れ値であるか否かを識別する．なお，関数 $\Phi(x)$ で表される累積標準正規分布表における 99%信頼区間は，$-2.58 \leq x \leq 2.58$ である．

$$f_1(C_0^{<1>}) = \begin{cases} 0 & if \quad m - 2.58 \times V \\ & \qquad < C_0 < m + 2.58 \times V \\ 1 & otherwise \end{cases} \tag{3}$$

ここでは正常値を 0，外れ値を 1 としている．

- 同じ曜日の同時刻における総トラフィック量
  総トラフィック量 $C^{<2>}$ に対して式 (1),(2),(3) を適用して外れ値を検出する．

- 同じ曜日の同時刻における平均トラフィック量
  平均トラフィック量 $C^{<3>}$ に対して式 (1),(2),(3) を適用して外れ値を検出する．

- 同じ曜日の同時刻における SSID 別接続端末種類数
  SSID ごとに求めた接続端末種類数を $C^{<4_j>}(j = 1, 2, ..., N_s)$ とする．ここで N_s は SSID の種類数である．$C^{<4_j>}$ に対して式 (1),(2),(3) を適用して外れ値を検出する．一つ以上の SSID で外れ値を検出した時を，本項目における外れ値とみなす．

- 全観測時間における一時刻前からの接続端末種類数の変動量
  注目している区間と一時刻前の区間での接続端末種類数の変動量の絶対値を $C^{<5>}$ とする．$C_0^{<5>}$ が過去 N_c 区間内の値と比較して大きい値である場合，すなわち変動量が大きい場合を外れ値とする．$C_k^{<5>}(k = 0, 1, ..., N_c)$ を降順にソートし，$C_0^{<5>}$ が上位 1%以内に含まれる場合に外れ値として検出する．ここで N_c は注目する区間の最大値である．

- 全観測時間における一時刻前からの総トラフィック量の変動量
  注目している区間と一時刻前の区間での総トラフィック量の変動量の絶対値を $C^{<6>}$ とする．$C_k^{<6>}(k = 0, 1, ..., N_c)$ を降順にソートし，$C_0^{<6>}$ が上位 5%以内に含まれる場合を外れ値として検出する．

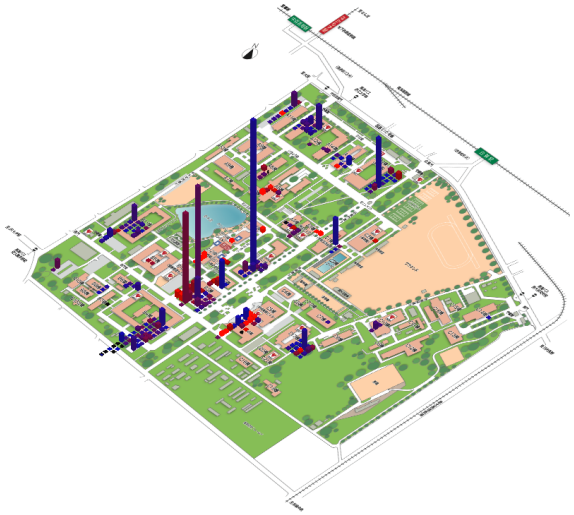

図 2: 可視化の例 (三次元描画)

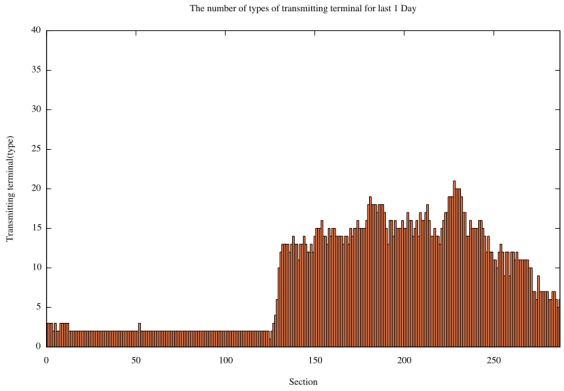

図 3: 可視化の例 (グラフ)

- 全観測時間における SSID 別の一時刻前からの接続端末種類数の変動量
  SSID ごとに求めた，注目している区間と一時刻前の区間での接続端末種類数の変動量の絶対値を $C^{<7_j>}$ とする．$C_k^{<7_j>}(k = 0, 1, ..., N_c)$ を降順にソートし，$C_0^{<7_j>}$ が上位 1%以内に含まれる場合を外れ値として検出する．一つ以上の SSID で外れ値を検出した時を，本項目における外れ値とみなす．

- 全観測時間における SSID 別の一時刻前からの接続端末種類数の変動率
  SSID ごとに求めた，注目している区間と一時刻前の区間での接続端末種類数の変動率の絶対値を $C^{<8_j>}$ とする．$C_k^{<8_j>}(k = 0, 1, ..., N_c)$ を降順にソートし，$C_0^{<8_j>}$ が上位 1%以内に含まれる場合を外れ値として検出する．一つ以上の SSID で外れ値を検出した時を，本項目における外れ値とみなす．

表 1: 異常発生日時および異常発生箇所

年月日	時刻	AP	年月日	時刻	AP
20170710	10:35	70	20170712	12:10	221
20170710	10:50	71	20170712	12:50	82
20170710	12:45	71	20170712	13:25	192
20170710	14:20	216	20170712	13:50	192
20170711	10:15	76	20170712	14:00	34
20170711	10:55	35	20170713	12:00	208
20170711	11:55	34	20170713	14:30	95
20170712	09:55	43	20170713	14:35	215
20170712	11:20	192	20170713	17:10	215

図 4: 2017 年 7 月 9 日における各 AP の接続端末種類数および総トラフィック量

3.4 可視化手法

あらかじめキャンパス内の AP をキャンパスマップにマッピングしておく．各 AP に対応した座標から特徴量の大きさに対応した色・高さの直方体を描画することで，抽出した特徴量を三次元的に描画する．直方体の色および高さは比率または絶対値によって表現することとし，比率で表現する場合は事前に色と高さに対する基準値を与えておくことで，その基準値に対する特徴量の割合を描画する．可視化の例を図 2 に示す．図 2 の例では，直方体の高さに総トラフィック量を与え，色は比率が 0.0 から 1.0 に近づくにつれ，青色から紫色を経て赤色となるようにしている．また 0.0 を特異点として扱い，黒色で表示している．比率が 1.0 以上の際は全て赤色で表示している．また，直方体をクリックすることで，その AP の特徴量の時系列情報を図 3 のようなグラフで表示する．このグラフは x 軸を区間（時刻）とし注目している時刻が右端（最大値）になるようにし，y 軸は管理者が指定した特徴量（図の例では，接続端末種類数）としている．このグラフにより，描画している時刻から一定の区間を遡った時刻までの指定した特徴量の変動を時系列順に確認することができる．

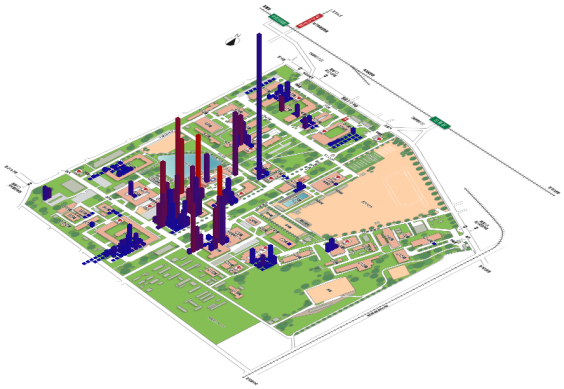

図 5: 2017 年 7 月 12 日における各 AP の接続端末種類数および総トラフィック量

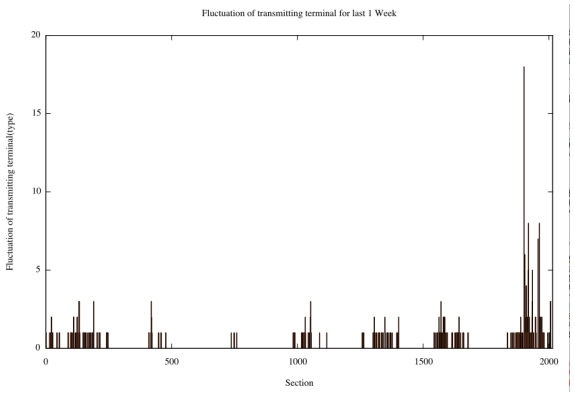

図 6: 215 番の AP における接続端末種類数の変動量の時系列変化

4. 実験

4.1 実験環境

大阪府立大学中百舌鳥キャンパスにおける 2016 年 12 月 14 日 0 時から 2017 年 7 月 13 日 23 時 59 分までの 7ヶ月分の全学無線 LAN 利用ログを対象に，異常状態の検出および可視化実験を行った．キャンパス内では 241 台の AP が稼動しており，無線 LAN コントローラにそれらの AP における無線 LAN 利用ログが蓄積されている．無線 LAN コントローラに単位時間ごとにアクセスすることで無線 LAN 利用ログを収集した．可視化に際し，あらかじめ本学中百舌鳥キャンパスのキャンパスマップ画像およびキャンパス内 241ヶ所の AP 設置場所とそれに対応するマップの座標は可視化システムに登録した．

4.2 実験と考察

全学無線 LAN 利用ログから特徴量を抽出し，2017 年 7 月 7 日 0 時から 7 月 13 日 23 時 59 分までの 1 週間分の全学無線 LAN 利用ログにおいて，3.3 節で述べた 8 つの判断基準で，外れ値を検出した基準が閾値以上あるかどうかを識別した．ここでは単位時間を 5 分，閾値を 7 と

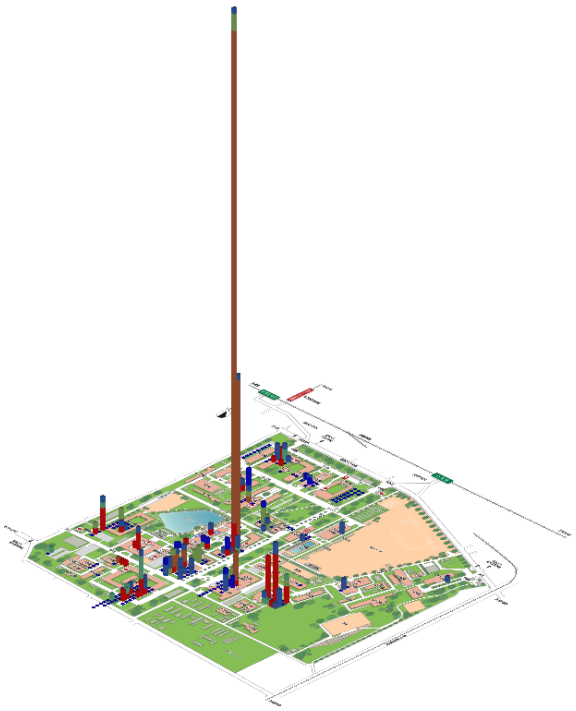

図 7: 2017 年 7 月 10 日 10 時 50 分における各 AP の SSID ごとの接続端末種類数

表 2: 異常の原因別の分類

休憩時間	カフェ	講義	不明	計
9	3	1	5	18

した．単位時間 5 分における一週間の区間数は 2016 であり，241 台の各 AP ごとに識別を行ったため，実験では 485,856 区間について実験を行った．実験の結果，異常と判断されたのは，18 区間であった．異常として判断された日時および AP の識別番号 (0,1,...,240) を表 1 に示す．表 1 より以下の 3 点がわかり，それぞれの点について調査する．

- 休日 (土曜日・日曜日) である 7 月 8 日および 7 月 9 日は，異常が検出されていない．

  休日と平日の AP ごとの接続端末種類数および総トラフィック量を図 4，図 5 に示す．図 4 では，休日（7 月 9 日）における本学中百舌鳥キャンパス全体の様子を可視化した．図 5 では，平日（7 月 12 日）の本学中百舌鳥キャンパス全体の様子を可視化した．なお，図 4，図 5 ともに直方体の高さに総トラフィック量を与え，色に接続端末種類数を与えている．ここでは，直方体の高さの基準値を 2.4GB，色の基準値を 100 台としている．図 4，図 5 より休日では平日に比べ本学の無線 LAN システムに繋がる端末が極端に少ないことがわかる．接続端末が極端に少ない場合，特徴量の多大な変動が観測されにくいため，異常が検出されていないと考えられる．

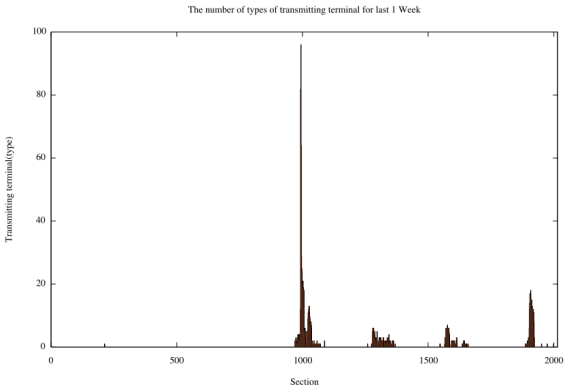

図 8: 71 番の AP における接続端末種類数の時系列変化

- 休憩時間に異常が多く検出されている．

  本学では 9:00〜10:30,10:40〜12:10,12:55〜14:25,14:35〜16:05,16:15〜17:45 の時間帯に講義を行っており，休憩時間およびその前後の時間帯である 8:50〜9:10,10:20〜10:50,12:00〜12:20,12:45〜13:05,14:15〜14:45.15:55〜16:25,17:35〜17:55 には，学生の教室への出入りが激しい．図 6 は，215 番の AP に着目し x 軸を 2017 年 7 月 7 日 0 時から 7 月 13 日 23 時 59 分までの 5 分ごとの区間とし，y 軸を接続端末種類数の変動量とするグラフである．図 6 より 1901 番の区間 (7 月 13 日 14 時 25 分〜29 分) では 18 台，1919 番の区間 (7 月 13 日 15 時 55 分〜59 分) では 8 台の変動があり，休憩時間およびその前後の時間では接続端末種類数の変動量が大きいことがわかる．したがって，休憩時間に異常が多く検出されている原因として，休憩時間における学生の流動に伴う AP の利用状況の変化が挙げられる．

- 192 番の AP において，異常が一週間の内に 3 回検出されている．

  192 番の AP は本学で唯一のカフェがある場所であり，講義時間・休憩時間に関係なく不規則に人が出入りするため，同曜日同時刻における特徴量を比較する判断基準では外れ値をとりやすいことが異常として検出された原因と考えられる．

検出された異常を可視化システムを用いそれぞれ検証する．ここでは，2017 年 7 月 10 日 10 時 50 分における 71 番の AP での異常を例に挙げる．可視化の例を図 7，図 8 に示す．図 7 は，2017 年 7 月 10 日 10 時 50 分における本学中百舌鳥キャンパス全体の様子を可視化した結果である．直方体の色は SSID ごとに対応するようにし，高さには SSID ごとの接続端末種類数を与えた．図 7 よ

りキャンパス中央部に周囲の直方体よりも一段と高い直方体がある．これは 71 番の AP を示す座標にあり，該当時刻に 71 番の AP で多数の端末が接続されていることを示す．この直方体は橙色が多くを占めているため，橙色に対応する SSID より，学生個人の所有する端末が多数を占めていることがわかる．また，異常とは検知されていないが，隣接する 70 番の AP でも多くの端末が接続されていることがわかる．図 8 は，71 番の AP に着目し，x 軸を 2017 年 7 月 7 日 0 時から 7 月 13 日 23 時 59 分までの 5 分ごとの区間とし，y 軸を接続端末種類数とするグラフである．図 8 より，993 番の区間 (7 月 10 日 10 時 45 分〜49 分) では 82 台，994 番の区間 (7 月 10 日 10 時 50 分〜54 分) では 96 台の端末が突発的に接続されていることがわかる．これは該当日時に大人数が受講する講義において受講生の所持するスマートフォンなどの端末を，一斉に全学無線 LAN システムに接続させ，情報システムに情報を入力させていたことを確認している．

その他の異常に関して，該当時刻における SSID の接続端末種類数を確認したところ，全ての異常において教職員が自由に使える SSID が多量に使用されており，異常発生時刻・場所および講義時間帯に影響されない点を考慮すると，これらの異常は教職員による会議に伴う利用状況の変化を検出したと考えられるが，今回用いたデータだけでは判断できないため，今後他の情報もあわせて確認することで異常を検出した原因を調査したいと考えている．異常として判断された 36 区間を原因別に分けた結果を表 2 に示す．

5. おわりに

本論文では，無線 LAN の利用状況を示すログを無線 LAN コントローラから取得し，ログから特徴量を抽出することで異常を検出し，それらの特徴量をマップに三次元的に描画して可視化することで管理者が異常を容易に把握できるように支援する手法を提案した．膨大な無線 LAN 利用ログデータを可視化することにより，管理者に対し時々刻々と変化する無線ネットワークの情報を視覚的に分かりやすく提示することができた．また，全学無線 LAN 利用ログに対し，本手法を適用することで，異常と考えられる事象を発見することができた．

今後の課題として，より長期間の無線 LAN 利用ログを収集し外れ値の検出精度を向上させることや，検出された異常のより詳細な検証を行うことが挙げられる．加えて，本研究では 8 項目の特徴量を用いて実験を行ったが，他に有意な特徴量があるかを十分に調査することも課題として挙げられる．

参考文献

[1] 鳩野 逸生, “全学無線 LAN 利用ログ情報の解析と応用,” 情処研報, Vol.2015-IOT-31, No.10, Vol.2015-SPT-15, No.10, 2015.

[2] 柳田 典章，廿日 出勇，青木 謙二，園田 誠，黒木 亘，川畑 圭一郎, “宮崎大学無線 LAN の利用状況の解析,” 情処研報，Vol.2015-IOT-31 No.9，Vol.2015-SPT-15 No.9，2015.

[3] 鈴木 和也，馬場 俊輔，和田 英彦, 中尾 康二, 高倉 弘喜, 岡部 寿男, “迅速な障害対応を支援するトラヒック可視化システムの構築と評価,” 信学論（B），Vol.J92-B，No.7，pp.1072-1083，2009.

[4] 鈴木 宏栄，衛藤 将史，井上 大介, “ネットワークトラフィック可視化システム NIRVANA の開発と評価,” 情報通信研究機構，Vol.57，pp63-80，2011．L-001，2016-09.

[5] “Aruba Networks 社: Access-Points,” http://www.arubanetworks.com/ja/products/networking/access-points/，2017 年 7 月 18 日 22:25.

[6] “Aruba 6.0 MIB,” http://gold.nvc.co.jp/supports/aruba/Relesenote/ArubaOS/ArubaOS6.0/User%20Guide/ArubaOS_60MG.pdf，2017 年 7 月 18 日 22:32.