

車両感知器データのマイニングによる 都市内の渋滞パターンの時空間分析

宮下 明久¹・井上 亮²

¹学生会員 東北大学博士前期課程 大学院情報科学研究科人間社会情報科学専攻
(〒980-8579 宮城県仙台市青葉区荒巻字青葉 6-6-06)

E-mail:miyashita@plan.civil.tohoku.ac.jp

²正会員 東北大学准教授 大学院情報科学研究科人間社会情報科学専攻
(〒980-8579 宮城県仙台市青葉区荒巻字青葉 6-6-06)

E-mail:rinoue@tohoku.ac.jp

都市内の道路交通は、季節・天候・曜日・時刻による交通需要の変動や、信号の制御パターン等の影響によって複雑な様相を示す。そのため、渋滞の発生時刻・場所、延伸・縮小過程、継続時間などの渋滞パターンを定量的に把握することは難しい。本研究は、情報工学分野で発達した「頻出パターンマイニング」を拡張し、一般道に設置された車両感知器データから、典型的な渋滞の時空間推移を効率的に抽出する手法を提案した。提案手法を沖縄県の車両感知器データに適用し、交通条件によって異なる渋滞パターンの抽出・分析を通して、その実行可能性と有用性を確認した。

Key Words : data mining, traffic sensor data, traffic congestion, spatio-temporal pattern

1. はじめに

道路交通渋滞は、所要時間増加による経済的損失や排気ガス・騒音による環境破壊など様々な問題を世界諸都市で引き起こしており、重要な都市問題の1つとして認識されている。これまで道路交通渋滞の軽減や解決に向け、渋滞発生をはじめとする道路交通状況の即時的な把握を目指した各種の取り組みが行われており、幹線道路を中心に設置した車両感知器を用いた常時観測もその一つである。車両感知器によって観測されたデータは、交通管理者による交通状況の監視や管制、道路交通利用者への情報提供に活用されている。

道路交通観測データの即時的利用に加え、長期間に渡る観測データの蓄積も行われている。蓄積された感知器データは、主に幹線道路上に限られた観測という点で空間的な網羅性には欠けるものの、都市内各所の道路交通状況が得られており、また、時間的には継続的かつ高頻度に観測されているため、渋滞発生など都市内の道路交通実態を把握する上で有用なデータと言える。

さて、都市内の道路交通は、季節や天候、曜日、時刻によって変動する需要に加え、多様な出発地・目的地を持つ様々な方向への移動、都市内に設置された信号の制御パターンなどの影響を受けるため、複雑な様相を示す。そのため、都市内における渋滞の発生時刻・発生地点や、

延伸・縮小過程、継続時間などで表される渋滞の典型的なパターンを、定量的に把握することは容易ではない。

これまで、渋滞分析の方法として、交通工学の交通流理論を応用したアプローチと、情報工学分野で発達した膨大なデータから有意な情報を取り出すデータマイニングを応用したアプローチが検討されてきた。前者では、ネットワーク上のボトルネックでの交通流の特徴に関する研究¹⁾²⁾や、Cell Transmission Modelを用いて実ネットワーク上の渋滞のボトルネックを発見する研究³⁾などが行われている。また、後者では、プローブデータや車両感知器データに対してデータマイニングを適用し、地域間⁴⁾⁵⁾や交差点間⁶⁾、リンク間⁷⁾の渋滞の時空間的な相関関係を発見する研究手法などが提案されている。

本研究は、後者のアプローチに着目し、大量のデータから一定頻度以上現れるパターンを抽出する手法である頻出パターンマイニング⁸⁾を基に、長期間に渡る継続的な観測で蓄積されてきた車両感知器データから、渋滞がいつ発生し、どこまで拡大するのかを表す「渋滞パターン」を抽出する都市内の渋滞実態把握手法の構築を行う。その後、沖縄本島で取得された感知器データを用いて、ボトルネックを先頭に発生した渋滞が周辺道路に影響を及ぼし拡大した後、縮小・解消に至る渋滞の典型的形態を、曜日や天候などの条件毎に抽出する分析を通して、提案手法の適用可能性や有用性の評価を行う。

2. 車両感知器データのマイニングによる渋滞把握に関する既存研究

車両感知器データに対して、データマイニング手法を活用し、渋滞の時空間遷移過程の分析を試みた先行研究としてLiu et al.(2015)⁷⁾が挙げられる。

この研究では、あるボトルネックを先頭に、渋滞が時間的・空間的に連続して拡大・縮小するパターンの抽出を目指した分析手法を提案している。渋滞リンクの時空間遷移過程の表現にはDynamic Bayesian Network (DBN) を採用し、DBNを用いて渋滞リンクの遷移関係を分析する、一連の渋滞が発生する空間領域を抽出するために、頻出パターンマイニングを使用している。

頻出パターンマイニングへの入力データは、車両感知器データを用いて各リンク・各時間帯の渋滞判定を行った上で、時空間的に連続な渋滞が発生しているリンク集合である。この入力データには、渋滞リンクの遷移を分析するのに必要な時間情報は含まれていない。

これまでにデータマイニング手法を活用し、渋滞が波及する空間領域を抽出する分析は行われてきたが、筆者らの知る限り、渋滞の時間的推移の抽出を目指した検討は行われていない。

3. 頻出パターンマイニング

(1) 頻出パターンマイニング

頻出パターンマイニングとは、大量のデータ内に一定頻度以上現れるパターンを抽出・列挙し、そのデータの特徴発見を目指す分析である⁸⁾。POS データを基に顧客の買物かご内の商品構成を分析する「バスケット分析」に端を発しており、パターンを構成する事象の最小単位をアイテム(商品)、一回の処理単位をトランザクション(取引)と呼び、一定数以上のトランザクションで観測されたアイテム集合を抽出する。

この分析では、「アイテム集合 X が発生する条件下 (条件部) で、アイテム集合 Y が発生 (結論部)」する共起関係を表す「相関ルール」を、「全トランザクションに対する、アイテム集合 $X \cup Y$ を含むトランザクションの割合」である支持度と、「アイテム集合 X を含むトランザクションに対する、アイテム集合 Y を含むトランザクションの割合」である確信度の 2 指標で評価する。支持度はアイテム集合 $X \cup Y$ の出現頻度を、確信度はアイテム集合 X とアイテム集合 Y の関連性の強さを表す。通常、最小支持度・最小確信度と呼ばれる閾値を設定して相関ルールを抽出し、アイテム集合 $X \cup Y$ を頻出パターンとして抽出する。

(2) FP-Growthアルゴリズム

頻出パターンマイニングの最速アルゴリズムの 1 つである FP-Growth アルゴリズム⁹⁾を概説する。これは、FP-Tree という木構造を用いて効率的に頻出パターンを列挙する、深さ優先型のアルゴリズムである。

なお、2. で示した先行研究では Apriori アルゴリズム⁸⁾を使用した分析を行っている。しかし、本研究では時間帯を考慮した渋滞リンクデータを用いた分析を検討しており、大規模の入力データにも適用可能な手法が求められる。そこで、本研究では、より効率的な頻出パターンマイニングアルゴリズムとして知られる FP-Growth を採用する。

a) FP-Tree

FP-Tree とは、トランザクションデータを格納する木構造である。データから最小支持度を満たさないアイテムを除き、また共通のアイテム集合を含む一部のトランザクションを集約して、頻出パターンの効率的探索に特化したコンパクトなデータを構築する。

ここでは、表-1 に示す 10 件のトランザクションから、最小支持度 20% として 2 件以上のトランザクションに含まれるアイテム集合を頻出パターンとして抽出する場合を例に、FP-Tree の構成手順を解説する。

手順①では、トランザクションに含まれるアイテムの出現頻度を調べ、その降順に並べたリストを作成 (図-1) する。アイテム d, f, g, h は最小支持度を満たさず、いかなる頻出パターンにも含まれる可能性がないため、以後の分析には必要ないことが確認できる。

次に、手順②で各トランザクションのアイテムを出現頻度の降順に並び替え、最小支持度を満たさないアイテムを除外した上で、図-1 の FP-Tree を作成する。各ノードには、アイテム・出現回数に加えて、親ノードおよび同一アイテムのデータが格納された次ノードへのポインタを属性として記録する。

作成される FP-Tree は頻出パターンの抽出に対する完全性を有しており、必要なすべてのデータが記録されている。例えば、アイテム e を含むトランザクションは、そのノードを指すポインタを辿ると、(b, a, e) が 1 件、(b, a, c, e) が 1 件、(a, c, e) が 2 件あることが分かる。なお、FP-Tree では、アイテム集合の共通部分の一部を集約し、代わりに頻度を記録して、データ量の圧縮を図っている。その圧縮効果を高めるため、通常、根ノードから出現頻度の高いアイテムを配置する。

表-1 トランザクションデータ(1)

番号	アイテム	番号	アイテム
1	b, c	6	a, c, e
2	b, d, f	7	b, c
3	a, b, e	8	a, c, e, g
4	a, b	9	a, b, c
5	a, b, c, e	10	h

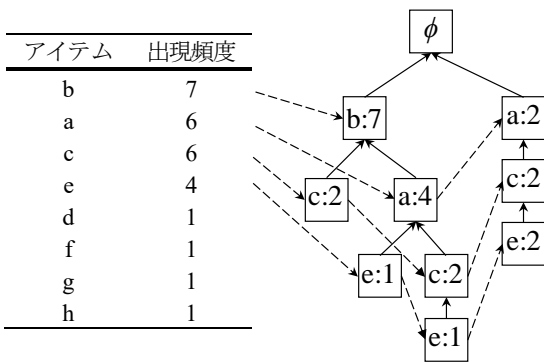


図-1 アイテムリストと FP-Tree

表-2 条件部{e}に関するトランザクション

アイテム	出現頻度
b, a	1
b, a, c	1
a, c	2

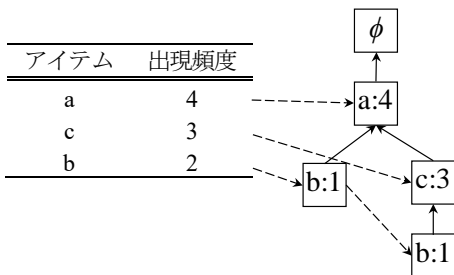


図-2 条件部 {e}に関するアイテムリストと FP-Tree

b) FP-Growthアルゴリズム

FP-Growth アルゴリズムは、手順①「FP-Tree のアイテムリストの下から順にアイテム 1 つを選んで条件部のアイテム集合に追加」、手順②「条件部のアイテム集合を含むトランザクションを FP-Tree から検索」、手順③「各アイテムの出現頻度を数える操作を通して頻出パターンを抽出」、手順④「手順②のトランザクションから条件付き FP-Tree を作成」の 4 手順を再帰的に繰り返す。

図-1 の例では、まず、アイテム e を条件部のアイテム集合に加えて {e} とする。次に、FP-Tree から条件部のアイテム集合 {e} を含むトランザクションを取り出し(表-2)、他のアイテムの出現頻度を確認する(図-2 アイテムリスト)。この例では、a, c, b のすべてが最小支持度を満たすため、{a, e}, {b, e}, {c, e} を頻出パターンとして追加する。また、表-2 のトランザクションから、条件部 {e} に関する条件付き FP-Tree を作成する。この操作を、図-2 アイテムリスト中のアイテムについて再帰的に繰り返すと、アイテム e を含むすべての頻出パターンを列挙できる。

なお、条件部へのアイテムの追加は、アイテムリストの降順、すなわち葉ノードの近くに記録されたアイテムから順番に行う。図-1 ではまずアイテム e を条件部に設定し、再帰的な操作を経てアイテム e を含むすべての

頻出パターンを抽出する。次に、アイテム c を含む頻出パターンを探索する場合には、アイテム e を含むパターンを考慮する必要がないため、トランザクションデータを取り出す際には、アイテム c を記録したノードから根ノードに至るパスだけを見ればよい。

以上の手順により FP-Growth アルゴリズムは、トランザクションの一部を集計して構築された FP-Tree を用いて、効率的に頻出パターンの列挙を行うことができる。

4. 頻出パターンマイニングアルゴリズムを拡張した渋滞パターン抽出手法の提案

本研究では、異なる日に、同じ時間帯に同じ道路リンクに渋滞が発生する状態を捉え、渋滞パターンとしての抽出を目指す。渋滞道路リンクとその渋滞観測時間帯からなる情報をアイテム、ある一日に出現するアイテムをトランザクションと考え、頻出パターンマイニングを適用すると、閾値として設定した日数以上で発生する渋滞パターンの抽出が可能である。しかし、その抽出結果の中には、渋滞道路リンクが空間的に繋がっていない、あるいは、渋滞時間帯が継続していないパターンも含まれることになる。本研究では、渋滞の発生・拡大・縮小・解消の過程の把握を目的としており、時間的・空間的に不連続なパターンは不必要な情報となる。不要なパターンの抽出は分析時間の増大を招き、分析の実行可能性の妨げになるため、効率的なパターン抽出が不可欠である。

本節では、効率的な渋滞パターン抽出に向けた、頻出パターンマイニングアルゴリズムに対する 2 つの拡張について説明する。

(1) 時空間的な接続関係の考慮

渋滞道路リンクの空間的接続関係や渋滞時間帯の時間的連続関係を考慮したパターンの探索範囲限定を通して、分析実行時間の短縮や実行可能性の向上を図るマイニングアルゴリズムについて示す。

前節で示した FP-Growth アルゴリズムでは、最小支持度を満たすアイテムを、条件部のアイテム集合に加えたアイテム集合を頻出パターンとして出力する。ここで、アイテム間の時空間接続関係を確認し、接続関係を有しないアイテム集合は頻出パターンとして抽出しない処理を行うと、効率的なパターン抽出が可能になる。

そこで、FP-Tree の構築過程や FP-Growth の実行過程を以下の通り修正し、アイテム間の時空間接続関係を考慮した渋滞パターン抽出方法を提案する。

まず、FP-Tree 構築の手順①は、条件部を満たしたトランザクションデータが入力された際に、アイテムの出現頻度に加えて、条件部のアイテム集合に含まれるアイ

テムとの接続関係を調べる。通常の FP-Tree 構築では、出現頻度の降順に並べたアイテムリストを作成するが、本研究では、接続関係を満たさないアイテム、接続関係を満たすアイテムそれぞれについて出現頻度の降順に並べたアイテムリストを作成する。次に、手順②では、各トランザクションのアイテムを、接続関係を満たさないアイテムリスト・満たすアイテムリストの順序に並び替え、最小支持度を満たさないアイテムを除外した上で FP-Tree を作成する。この FP-Tree では、根ノードに近いノードに条件部のアイテム集合と接続関係を持たないアイテムの情報が、葉ノードに近いノードに接続しているアイテムの情報が記録される。

この FP-Tree に対して、FP-Growth の手順①を、FP-Tree の接続関係を有するアイテムリストの下から順にアイテム 1 つを選んで条件部のアイテム集合に追加し、そのアイテム集合を頻出パターンとして抽出するよう変更する。

この手順では、時空間接続関係を満たすアイテム集合のみを条件部に設定して分析するため、FP-Growth よりも分析範囲を限定した効率的な分析が可能である。

(2) 頻出アイテム集合の記号化による飽和集合探索の効率化

都市では、ほぼ同時刻に同一ボトルネックを先頭に発生した渋滞が長時間継続し、周辺道路に影響が波及する様子が毎日のように観察される。そのため、ある時間帯に常に渋滞する接続する道路リンクの組み合わせや、ある道路リンクが常に渋滞する時間帯など、渋滞データには非常に高い割合で観察されるアイテム集合が存在する。

長時間継続し広範囲に広がる渋滞の分析では、多くのアイテムが頻出渋滞パターンを構成する。頻出パターンマイニングアルゴリズムは、分析過程で頻出パターンのすべての部分集合を列挙するが、あるパターンの部分集合の数はアイテム数に対して指数関数的に増加するため、分析の実行可能性が大幅に低下してしまう。

この問題への対処として、渋滞データの特徴である互いに同時に出現する割合が高いアイテム集合（共出現集合）を、頻出パターンマイニングの適用前に集約・記号化して対応する。この対応により、入力データ量も減少し、かつ、集約した共出現集合の部分集合は列挙されなくなるため、飽和集合以外のパターン出力が大幅に減少し、計算時間の短縮効果が期待できる。

本研究では、共出現集合の評価基準として FP-Growth アルゴリズムでは考慮していない「確信度」を援用する。条件部をアイテム集合を構成する各アイテム、結論部を共出現集合とした確信度を求め、その最小値を共出現集合の判定基準として、「記号化基準」を超える場合にそのアイテム集合を記号化する。

記号化基準によって記号化が行われる共出現集合が変

表-3 トランザクションデータ(2)

番号	アイテム
1	a, b, c
2	a, b, c
3	a, b, c, d
4	a, b, c, d
5	a, b, d

化するため、抽出結果は記号化基準の値に影響を受ける。ここで、5 件のトランザクションからなるデータ(表-3)から、最小支持度 60%のアイテム集合を頻出パターンとして抽出する場合を例に、記号化基準値の変化による抽出結果に対する影響を解説する。

記号化を行わない場合には、すべての部分集合を含む頻出パターン {a}, {b}, {c}, {d}, {a, b}, {a, c}, {a, d}, {b, c}, {b, d}, {a, b, c}, {a, b, d} が列挙される。このうち飽和集合は {a, b, c}, {a, b, d} の 2 つである。

記号化基準を 100%に設定すると、共出現集合 {a, b} が記号化され、抽出される頻出パターンが {a, b}, {a, b, c}, {a, b, d} の 3 つに減少するが、2 つの飽和集合はともに抽出される。しかし、記号化基準を 80%に下げると、共出現集合 {a, b, c} を記号化するため、{a, b, c} のみが抽出され、他方の飽和集合 {a, b, d} は抽出されなくなる。

記号化基準を 100%に設定した場合は、記号化しない場合と同じ飽和集合を列挙できる。しかし、記号化基準を下げると、多くのアイテム集合を集約して、入力するアイテム数を減らし、計算の高速化やアルゴリズムの実行可能性の向上が可能だが、記号化したアイテム集合の一部が欠けた飽和集合を抽出できない。このように記号化基準を下げることで抽出結果の正確さはトレードオフの関係にあるため、記号化基準はできる限り 100%に近い値で実行することが望ましい。

以上の2種類の拡張により、探索範囲を限定し、部分集合の抽出を減らすことができるため、本研究で使用するアルゴリズムは、計算時間の短縮など、大規模データに対する適用可能性を向上させることができる。

5. 沖縄本島南部の車両感知器データへの適用

(1) データ概要と渋滞リンク情報の作成

沖縄県警より提供された沖縄本島内の車両感知器による観測データを用いる。2011年6月1日から2013年9月30日までの853日間、沖縄本島南部の二次メッシュ（那覇392725, 糸満392715, 知念392716, 与那原392726）に属するVICSリンク445本における観測結果で、5分間隔の交通量と時間平均速度が記録されている。

各リンクで観測された車両感知器データから臨界速度を推定し、それを基準に各リンク・時間帯の渋滞の有無

を判定して渋滞リンク情報を作成する。なお、曜日による車種・運転者層の違いや気象条件の違いが渋滞発生に与える影響を分析するため、平日の各曜日・土休日別、無降雨時・荒天時別に区分し臨界速度推定を行う。なお、気象条件は、気象庁沖縄気象台那覇観測所のアメダス・時間雨量データを使用し、0mmを無降雨時、5mm以上を荒天時と設定した。全期間20,472時間のうち、無降雨時18,831時間(約92%)、荒天時は378時間(約1.8%)である。

ここで、車両感知器データの例として、平日・無降雨時の国道58号線下り・旭橋交差点(図-3)における基本図を示す。なお、旭橋交差点はボトルネック交差点である。整数値で記録されている交通量・平均速度から例外的な観測結果を除外するため、交通量1台以上の総観測時間帯数の0.01%以上の頻度で観測された交通量・平均速度の組み合わせのみを以後の分析に用いる。なお、図-3では、0.01%以上の頻度があるデータ点を赤、より頻度が小さいデータ点を寒色系に着色している。また、交通密度は0.5台/km刻みの値として求める。交通密度の推定には空間平均速度が必要だが、車両感知器データの時間平均速度で代用している。そのため、平均速度が過大で、交通密度を過小に推定していることには注意を要する。

次に、交通密度と平均速度の関係を線形式で推定し、臨界速度を推定する。K-V図上で、交通密度が低い領域では平均速度が極端に高く、交通密度が高い領域では平均速度が著しく低い傾向が見られた。この状況は当該リンクの性能評価上は参考にならないと考えられるため、臨界速度推定から除外するのが適当であろう。そこで、交通密度5台/km以上、平均速度20km/h以上のデータ点のみを選んで、各交通密度における最大平均速度を抽出した後、平均速度を被説明変数、交通密度を説明変数とする線形回帰を行う。なお、図-3のK-V図上の黒線は、交通密度と平均速度の関係の推定結果を示す。

さて、平均速度と交通密度が線形関係だと仮定すると、交通量最大となる臨界速度はK-V図の切片の1/2に相当する。この推定値をQ-V図上の赤太線で表現すると、図-3(b)では、臨界速度は交通量最大の平均速度とほぼ一致しており、推定の妥当性が確認できる。

なお、各リンクの渋滞判定は5分間隔で行っているが、車両感知器の観測誤差や、信号制御のサイクルなどの影響により、渋滞判定結果は安定しない。そこで本研究では、分析間隔を15分とし、その間に1回以上渋滞があると判定された場合に渋滞リンク情報を設定する。

(2) 提案手法の実行可能性の検証

(1)で作成した渋滞リンク情報を用いて、提案手法の実行可能性を検証する。まず、提案手法の前処理における、頻出集合の記号化による本処理への影響を確認する。なお、Intel(R) Xeon(R) 2.8GHz・メモリ 8GBを搭載した

Windows10のPCを利用した。プログラムはC++で作成し、Visual Studio 2015を用いてコンパイルを行った。

平日の中で渋滞リンク数最小の月曜日(図-4)について、夕方16時から20時を対象に、記号化基準を100%で共出現集合を記号化する。また、気象庁沖縄気象台那覇観測所のアメダス・時間雨量データを使用して、分析時間帯に1時間当たりの降水量が0.1mm以上の時間帯が含まれていない場合を「晴天」と定義し分析対象にした。なお、分析時間帯に1時間当たりの降水量が0.1mm以上の時間帯が含まれている場合を「雨天」と定義する。

このとき、共出現集合の記号化の有無と頻出パターンマイニングの実行時間・抽出パターン数と最小支持度の関係を図-5に示す。記号化なしの場合は、記号化ありの場合に対し、最小支持度が90%前後では実行時間に約100倍、抽出パターン数に約300倍の差が生まれている。

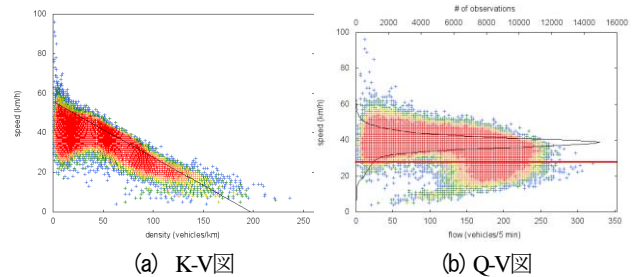


図-3 国道58号線下り旭橋交差点平日・無降雨時(2次メッシュコード392725, VICSリンク番号6)

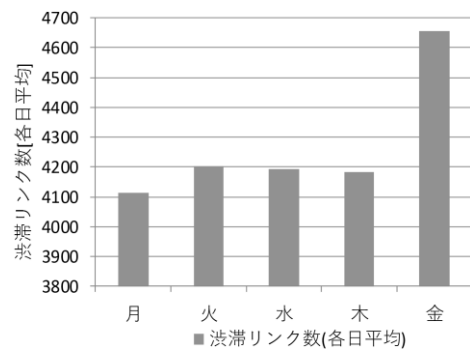


図-4 各曜日の渋滞リンク数平均値

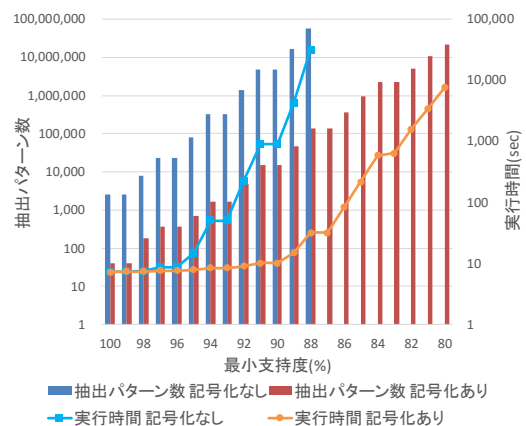


図-5 最小支持度と実行時間・抽出パターン数の関係

表-4 実行時間・飽和集合数

		実行時間(秒)	飽和集合数
晴れ	金曜日	3390.4	9,450
	月曜日	61.1	3,140
雨	平日	3391.1	15,067

表-5 各天候条件を満たす日数

	月	火	水	木	金	平日
晴れ	69	78	80	86	79	392
雨	12	17	17	9	15	70

表-6 那覇市中心部選定パターン

		天候	曜日	延ベ リンク数	最大時 リンク数
パターン①	晴れ		金曜日	72	9
			月曜日	45	9
	雨	平日	67	10	
パターン②	晴れ		金曜日	68	9
	雨		平日	75	9
パターン③	晴れ		金曜日	52	6
			月曜日	34	4
	雨		平日	42	6

閾値を下げると実行時間と抽出パターン数は指数的に伸びるが、明らかに前処理ありの場合がパターン数と実行時間ともに抑制できていることが確認できる。

以上より、前処理で記号化を行うことで、より実行可能性を確保できることが示された。

(3) 提案手法の有用性の検証

天候や曜日などの条件毎の渋滞パターンを抽出・比較し、本研究の提案手法の有用性を確認する。

渋滞リンク数最大の金曜日(図-4)でも最小支持度60%の分析が行えるように、記号化基準を90%と設定した場合の分析結果の概要を表-4に示す。天候が晴れの場合は、渋滞リンク数最大の金曜日と最小の月曜日の結果を、天候が雨の場合は、曜日毎では雨天日数が少ない(表-5)ため、平日全てのデータを集約して分析した結果を示す。

抽出されたパターンから、沖縄本島南部の中心都市である那覇市中心部周辺の時空間的に大きい渋滞パターンを各条件の分析結果から選定する。渋滞の時空間過程に着目した分析を行うため、16時から20時までの全時間帯で継続的に渋滞するパターンから、全時間帯の延べ渋滞リンク数が最大で、かつ、渋滞リンク数最大時の構成リンク数最大のパターンから順に、重複が無いパターンを選定した。結果を、表-6と図-6に示す。なお、表-6と図-6において同じパターンを同一の色で表示している。

図-6より、各渋滞パターンが、(a)の時間帯から渋滞

のピークを迎える(c)、(d)の時間帯に向けて拡大し、その後は縮小している様子が確認できる。また、同じ曜日・天候条件内でも、パターン①の方がその他の渋滞パターンと比べて渋滞が拡大する時間帯が早く、各渋滞パターンの拡大・縮小する時間帯が異なることが分かる。さらに、各時間帯において渋滞パターンは空間的につながることがなく、それぞれが独立した渋滞パターンであることが確認できる。

次に、晴天時の曜日による渋滞パターンの違いに着目する。表-6に示す渋滞が最も拡大した時のリンク数、パターン全体を構成している延べリンク数、図-6に示す渋滞拡大範囲から、金曜日は月曜日に比べ渋滞がより長時間、広範囲で継続していることが確認できる。

一方、天候の違いに着目すると、雨天時の渋滞パターンは、晴天時の金曜日の渋滞パターンと類似し、月曜日の渋滞パターンよりも、空間的に広範囲に拡大し、周辺の道路にまで影響を与える様子が確認できる。また、渋滞パターンの構成リンク数最大となる時間帯は(c)の前後であり、晴天時よりも早いことが分かる。このようなパターンが抽出された理由としては、雨天時は国道などの主要な幹線道路沿いで交通量の増加や処理能力の低下が起き、渋滞の範囲が早い時間帯から拡大したためであると推察される。

以上の3条件下での渋滞パターン抽出結果の考察から、提案手法は、曜日・天候などの条件の違いを明確に表す渋滞パターンを抽出できることが明らかになった。また、約2年3か月間の車両感知器データから、16時から20時まで夕方ラッシュ時の比較的長い時間帯を分析対象とした場合も、時空間的に連続な渋滞パターンを抽出することができることが示され、渋滞パターン分析に対する本提案手法の有用性が確認できた。

6. 結論

本研究は、車両感知器による交通観測データから日常的に発生する渋滞パターンを抽出する方法として頻出パターンマイニングに着目し、時空間的に連続な渋滞パターンを効率的に抽出できるアルゴリズムを構築した。沖縄本島南部の感知器データに対する適用を通して、提案手法は、曜日や天気・時間帯などの分析条件の異なる渋滞パターン抽出結果の比較から、渋滞の時空間的様相の違いを把握できることを確認した。

なお、今後の課題として、以下の項目が挙げられる。

本研究では、渋滞パターンの時間的変化を観察するため、渋滞の継続時間が長く、最大渋滞リンク数が多い渋滞パターンに着目して分析を行った。しかし、着目する観点を変えると、同一のボトルネックに起因する渋滞パ

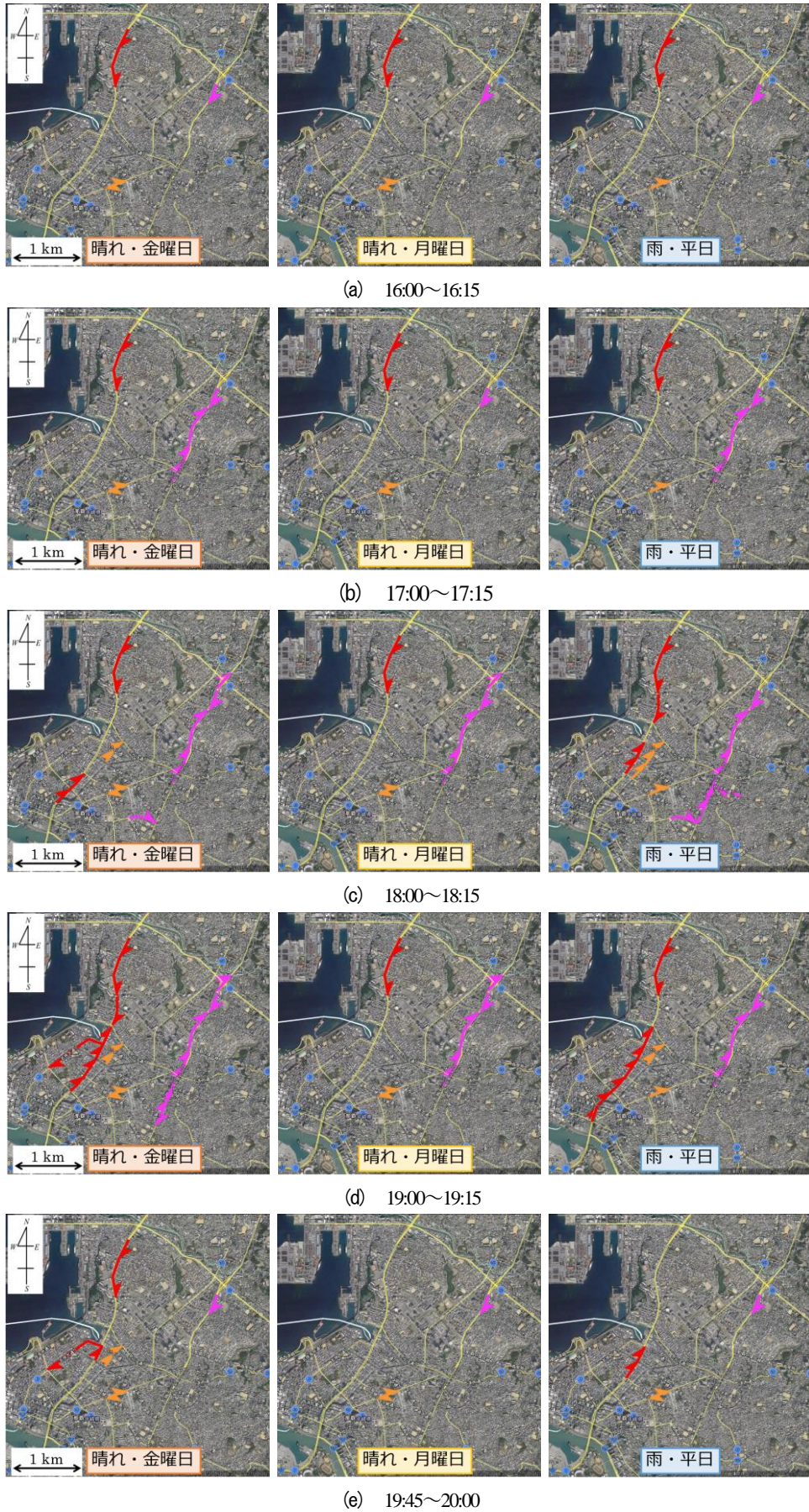


図-6 那覇市中心部の渋滞パターンの変化する時間帯

ターンでも、渋滞の継続時間は短い最大渋滞リンク数が極端に多いパターンが存在し、それによって異なる分析結果や考察が得られる可能性がある。都市部の渋滞現象を把握するためには、より多くの抽出結果を基にした分析方法を検討する必要がある。

また、前処理で高頻度に共出現する渋滞リンクを記号化する方法についても検討を要する。記号化基準を下げると分析手法の実行可能性は確保できるが、記号化された共出現集合のリンクの一部を含む、その他の渋滞パターンが抽出されなくなる問題が存在する。記号化基準を下げずに、低い最小支持度設定で分析できる、効率的な分析手法の構築が求められる。

また、本研究では、沖縄本島南部を対象に分析を行ったが、他都市でも同様に車両感知器データが蓄積されている。今後は他都市のデータを用いて、より大規模なデータへの適用可能性、各都市の特徴的渋滞パターンの抽出についても検証を行う必要があると考えている。

謝辞：本研究は公益財団法人日本道路交通情報センターとの共同研究の成果である。また、本研究は、JSPS科研費 26630233および15H04053の助成を受けたものである。

参考文献

- 1) Newell, G. F. 1998. A moving bottleneck. *Transportation Research Part B: Methodological*, Vol. 32, No. 8, pp. 531-537.
- 2) Ni, D., and Leonard, J. D. 2005. A simplified kinematic wave model at a merge bottleneck. *Applied mathematical modelling*, Vol. 29, No. 11, pp. 1054-1072.
- 3) Long, J., Gao, Z., Ren, H., and Lian, A. 2008. Urban traffic congestion propagation and bottleneck identification. *Science in China Series F: Information Sciences*, Vol. 51, No. 7, pp. 948-964.
- 4) Liu, W., Zheng, Y., Chawla, S., Yuan, J., and Xing, X. 2011. Discovering spatio-temporal causal interactions in traffic data streams. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, United States of America, August 21-24, 2011*, pp. 1010-1018.
- 5) Pang, L. X., Chawla, S., Liu, W., and Zheng, Y. 2013. On detection of emerging anomalous traffic patterns using gps data. *Data & Knowledge Engineering*, Vol. 87, pp. 357-373.
- 6) Chu, V. W., Wong, R. K., Liu, W., and Chen, F. 2014. Causal structure discovery for spatio-temporal data. *Proceedings of the 19th International Conference on Database Systems for Advanced Applications, Bali, Indonesia, April 21-24, 2014*, pp. 236-250.
- 7) Liu, W., Nguyen, H., and Chen, F. 2015. Discovering congestion propagation patterns in spatio-temporal traffic data. *Proceedings of the 4th International Workshop on Urban Computing, Sydney, Australia, August 10, 2015*.
- 8) Agrawal, R., and Srikant, R. 1994. Fast algorithms for mining association rules in large databases. *Proceedings of the 20th International Conference on Very Large Databases, Santiago de Chile, Chile, September 12-15, 1994*, pp. 487-499.
- 9) Han, J., Pei, J., Yin, Y. and Mao, R. 2004. Mining frequent patterns without candidate generation: a frequent-pattern tree approach. *Data Mining and Knowledge Discovery*, Vol. 8, pp. 53-87.

MINING SPATIO-TEMPORAL PATTERNS OF CONGESTED TRAFFIC IN URBAN AREA FROM TRAFFIC SENSOR DATA

Akihisa MIYASHITA and Ryo INOUE