

ETC2.0 プローブデータを活用した 豪雨災害時の交通ネットワーク異常検知

稲場 亘¹・柳沼 秀樹²・寺部 慎太郎³・田中 皓介⁴

¹学生非会員 東京理科大学大学院 理工学研究科土木工学専攻修士課程

(〒278-8510 千葉県野田市山崎2641) E-mail:7619503@ed.tus.ac.jp

²正会員 東京理科大学 理工学部土木工学科 (〒278-8510 千葉県野田市山崎2641)

E-mail: yaginuma@rs.tus.ac.jp

³正会員 東京理科大学 理工学部土木工学科 (〒278-8510 千葉県野田市山崎2641)

E-mail: terabe@rs.tus.ac.jp

⁴正会員 東京理科大学 理工学部土木工学科 (〒278-8510 千葉県野田市山崎2641)

E-mail: tanaka.k@rs.tus.ac.jp

災害時には道路リンクの途絶により、広域的な混雑や局所的な孤立が発生するため、平常時と比較して人々の交通行動パターンが変化すると考えられる。近年では、ETC2.0プローブデータを活用した道路交通状態のモニタリングが実施されており、災害時の通行可能区間に関する情報提供などが行われているが、発災による異常な交通状態を事前に把握するには至っていない状況にある。本研究では、災害による異常な交通状態の早期把握を念頭にETC2.0プローブデータを活用した交通状態の異常検知に取り組む。具体的には、原ら(2016)¹⁾が提案した道路道路ネットワークの疎構造学習に基づく状態推定を援用して、蓄積データと当該時刻の入力データが持つ相関構造の変化から異常検知を行い、災害等による異常な交通状態を自動的に抽出・把握することを目指す。本論では、2015年9月の関東・東北豪雨で被災した茨城県常総市周辺のデータを用いて提案を適用し、災害時の交通状態把握に貢献することを示す。

Key Words : *Anomaly detection, Heavy rain disaster, ETC2.0 probe data, Graphical Lasso*

1. はじめに

近年、全国各地で大規模な水害の発生が後を絶たない。今年も熊本県を中心に九州や中部地方を襲った令和2年7月豪雨では多数の人的、経済的な被害が生じた。水害発生時には住民が迅速に避難することが求められるが、浸水等で道路が途絶し、局所的な孤立や広域的な混雑が発生して大きな障害となっている。災害時に行政が避難誘導や道路啓開を効率的に実施するためには、道路ネットワークの異常を早期に把握できることが望ましい。

現在、道路ネットワーク上の交通状態は、車両感知器やプローブカー等により、部分的にモニタリングされており、事後的に通行不可能箇所の特定制などが行われているが、十分に機能しているとは言い難いのが現状である。そのような中、国土交通省は次世代型ETCであるETC2.0の普及を進めており、ETC2.0を搭載した車両から得られるプローブデータ(ETC2.0プローブデータ)により、全国を対象に車両の走行経路や速度などの詳細な移動情報

が収集可能となっている。ETC2.0の普及率は現状では2割程度²⁾に留まるが、将来的に普及の促進が見込まれるため、災害時の早期実態把握に有効となることが期待される。また、ETC2.0は常時データを取得しているため、平常時のデータを活用できるという強みがある。

以上を踏まえて、本研究では平常時に蓄積されETC2.0プローブデータを活用して、災害時に観測される異常データの検出モデルを構築する。具体的には、原ら¹⁾が提案した道路リンク間の相関構造をガウシアングラフィカルモデルによって記述する交通状態推定モデルを援用し、平常時における道路ネットワークの疎構造学習を行う。その上で、部分的に観測された災害時のデータについて変数間の相関の崩れを表す異常度を計算することで、災害時の異常データを検出することを試みる。ケーススタディとして、2015年9月の関東・東北豪雨で被災した茨城県常総市周辺における1ヶ月間のETC2.0プローブデータを用いて上記のモデルを実装し、交通状態把握の有効性について検証する。

2. 既往研究の整理と本研究の位置づけ

(1) プローブデータを活用した災害時の交通把握に関する研究の整理

車両プローブデータを活用して災害時の救援活動や道路啓開等の支援を目指す研究の多くは、事後的な状況分析を対象としている。例えば、秦ら³⁾は民間プローブ情報から道路状況の把握を試みている。具体的には、中越沖地震後の新潟県を対象に分析を行い、一定の限界はあるものの、広域的に道路の通行可能状況を判断できる情報が生成可能であることを示している。ETC2.0プローブデータを活用した事例としては、水谷ら⁴⁾が、ETCプローブ情報および民間プローブ情報を利用して、災害時における車両の通行経路の把握が可能であることを明らかにしている。ここでは、大雨災害時を対象に、通行止めとなった道路に並行する道路の通行実績から、代替路として利用された道路が確認できることを示している。また、ETCプローブ情報と民間プローブ情報の比較を行い、両者のメリット、デメリットを整理している。

(2) 交通状態推定と異常検知に関する研究の整理

プローブデータを用いて交通状態推定を試みる研究には、原ら⁵⁾があり、データ欠損リンクの空間的補間を行っている。過去の蓄積されたプローブデータから、道路ネットワークを構成する各道路リンク間の相関関係の疎構造学習を行い、未観測リンク速度を同時刻の観測リンク速度から補間することで道路ネットワーク全体の交通状態推定を行う手法を提案した。タイのバンコク中心部のタクシープローブデータを用いたケーススタディ分析では、未観測リンクが従来の手法よりも高い精度で補間可能であり、リンク間の局所的な近接関係だけでなく、大域的な相関関係を捉えることに成功している。

一方、プローブデータを活用して交通異常を検知する研究はETC2.0が普及する以前から存在するが、災害時を対象としたものは多くはない。例えば丸橋ら⁶⁾は、テンソル分解を用いて広範囲のセンサーから観測された様々な交通情報を統合し、道路全体の状態を把握するための手法を開発している。ETC2.0を活用した事例としては、増田ら⁷⁾が、高速道路内における渋滞先頭における渋滞速度回復距離により事故や路上障害物等の交通を遮断する可能性がある事象を検知するアルゴリズムを開発している。

(3) 本研究の位置づけ

プローブデータには時間的・空間的な未観測リンクが多く存在するため、観測データの挙動のみで交通状態把握に活用することは困難である。そこで筆者ら⁷⁾は、既往研究⁵⁾の手法を本研究と同一の水害被災地域に適用し、

全体的な傾向としては一定の精度で状態推定が可能であることを示した。しかしながら、災害時特有の局所的な渋滞の予測が難しかったことから、平常時と比較して災害時には道路ネットワーク構造が部分的に変化していることが示唆された。

そこで、本研究では災害時に道路ネットワークの相関構造が崩れることを利用して、異常な観測データを検知することを目指す。具体的には、浸水や避難交通の発生などにより、平常時とは異なるODが生じて相関構造が変化する兆候を検出することを試みる。適用上の課題として、未観測リンクの異常は検出できない点が挙げられる。すなわち、浸水等で通行不可能であったリンクについては明らかに異常と判定されるべきであるが、データが観測されない理由として、道路被害とは関係なく単にプローブカーが当該道路を走行しなかった場合も想定されるため、両者を区別することは難しい。第4章の基礎分析で平常時のリンク欠損状況から、災害時に通行不可能であったリンクの推定を試みるが、現段階の異常値検出モデルの枠組みでは対象外としている。このように、災害時の実現象を表現する上で問題は存在するが、部分的な観測データから災害発生時の交通異常に関する徴候を自動検出することで、災害時における交通状態把握に関する有効性を検証することを目的としたい。

3. 未観測道路リンクの空間補間モデルの構築

本研究では、原ら⁵⁾の交通状態推定モデルを援用する。以下にモデルの概要を記述するが、表記等は原著論文¹⁾および筆者ら⁷⁾を踏襲している。

(1) モデルの仮定

道路ネットワーク上の各方向別道路リンクが交通状態をもつと仮定し、各方向別道路リンクの集合を V とする。このとき、各時間単位において、道路リンクは交通状態 x_i をもつ。ここで、 $x = (x_1, \dots, x_{|V|})^T$ であり、本研究では交通状態 x_i を1時間単位の平均リンク速度とする。

同一時間帯の交通状態データは、ただ一つの多次元正規分布から生成されていると仮定する。

$$x \sim N(x|\mu, \Sigma) \quad (1)$$

ここで、 $N(\cdot)$ は多次元正規分布を表し、 μ と Σ は正規分布のパラメータである平均ベクトルと分散共分散行列である。なお、交通状態 x の生成モデルに単峰性を仮定しており、データはこれを満たすことに注意したい。

(2) 多次元正規分布のパラメータ推定

実データから多次元正規分布のパラメータ μ, Σ (または精度行列 θ)を推定する方法を示す。なお、精度行列 θ は分散共分散行列 Σ の逆行列であり、以降は計算の都合により Σ の代わりに θ を用いる。

まず、一般的なアプローチとしてデータから最尤推定値を求める方法が挙げられる。多次元正規分布の対数尤度関数は、

$$\log p(\mu, \theta) = \frac{|D|}{2} \log \det \theta - \frac{1}{2} \sum_{d \in D} (x^d - \mu)^T \theta (x^d - \mu) + \text{const} \quad (2)$$

で与えられる。ここで、 D はデータの集合、 x^d は各データの交通状態である。対数尤度関数をパラメータ μ および θ で偏微分することにより、 μ の最尤推定値は標本平均 m 、 θ の最尤推定値は標本分散共分散行列 S の逆行列として求めることができる。しかしながら、この方法ではデータ数に対してパラメータ数が非常に大きい場合に、 S が必ずしも最尤推定量とはならないことや、パラメータの過学習に陥ることが知られている⁸⁾。本研究が扱うモデルは、パラメータ次元がデータに比べて大きな問題であり、道路リンク数が1000であれば最低でも約50万個のデータ数が必要となる。そのため、本研究でこのアプローチを採用することはできない。そこで、正則化によって過学習を抑制する手法を適用する。このアプローチは、原らりが採用したアプローチであり、本研究では概説のみ行う。

正則化とは、パラメータ数に対してデータ数が不足している等の不良設定問題を対象に、もとの目的関数に罰則項を加えることで過学習を防ぐアプローチである。特に、L1ノルム正則化項を加えて最適化問題を解くことをラッソと呼び、多くのパラメータが0となることが知られている。

本研究における正則化手法の定式化を行う。道路ネットワーク上の交通状態 x を表現する多次元正規分布の確率モデルをマルコフ確率場としてみた場合、そのグラフをガウシアングラフィカルモデル（以下GGM）と呼ぶ。

$$p(x) = \frac{|\theta|^{1/2}}{(2\pi)^{D/2}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \theta (x - \mu) \right\} \quad (3)$$

このとき、任意のリンク $i, j \in V$ に対して、

$$p(x_i, x_j | x_{V \setminus \{i, j\}}) \propto \exp \left(\frac{1}{2} (\theta_{ii} x_i^2 + \theta_{ij} x_i x_j + \theta_{jj} x_j^2) + A x_i + B x_j + C \right) \quad (4)$$

と書くことができる。ここで、 A, B, C は x_i, x_j に依存しな

い定数であり、確率密度関数は積に分解できる。GGMの構造を学習することが、パラメータ θ を推定することに対応しており、疎なGGMの構造学習を可能にする手法がグラフィカルラッソ（以下GL）^{9),10)}である。本研究では、道路リンク間の共分散構造の推定にGLを用いる。

式(2)で示した多次元正規分布の対数尤度関数は、標本分散共分散行列 S と精度行列 θ を用いると次のように書き表すことができる。

$$\log p(\mu, \theta) = \frac{|D|}{2} (\log \det \theta - \text{tr}(S\theta)) + \text{const} \quad (5)$$

ここで、 $\text{tr}(\cdot)$ はトレースを表す。この対数尤度関数に対してL1ノルム正則化項を加えた最適化問題を解き、精度行列のパラメータを推定する。

$$\theta^* = \arg \max_{\theta} (\log \det \theta - \text{tr}(S\theta) - \rho \|\theta\|_1) \quad (6)$$

ここで、 $\|\theta\|_1 = \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} |\theta_{ij}|$ は θ のL1ノルムを表す。

L1ノルムは微分不可能であるため、式(6)は解析的に解くことができない。そこで、ブロック座標降下法を用いて求解する。ブロック座標降下法は、変数をブロックに分割し、あるブロックを変数とし、残りのブロックを定数と見なした問題を逐次的に解いて最適化を行う手法である。この方法により、制約付き最適化問題をL1正則化付き回帰問題に帰着させてパラメータの推定値を得ることができる。なお、GLは完全データに対して適用することを前提としているため、本研究のような欠損を含むデータに対してそのまま適用することはできない。そこで、各リンク速度の平均値を代入して欠損値を補間する前処理を行う。こうしてできた完全データのデータセットに対してGLを適用することにより、分散共分散行列の構造を推定する。

(3) 異常度の計算方法と外れ値検出手法

上記(2)で実データから推定された多次元正規分布のパラメータ θ を用いて、観測された各変数の異常度を計算する手法について述べる。

情報理論においては、情報量は確率分布の負の対数尤度と結び付けて定義され、観測値 x' に対する異常度 $a(x')$ は以下のように表現できる。

$$a(x') = -\ln p(x'|D) \quad (7)$$

式(7)は単調減少関数であるため、出現確率が小さい観測値は異常度が高く、出現確率が大きい観測値は異常度が低いという性質を満たす。GGMの範囲では、条件付き分布を計算することにより、任意のリンク i の異常度を以下のように定義することができる。

$$a_i(x') = \frac{1}{2} \ln \frac{2\pi}{\theta_{i,i}} + \frac{1}{2\theta_{i,i}} \left(\sum_{j=1}^D \theta_{i,j} x'_j \right)^2 \quad (8)$$

式(8)で各時間帯における観測リンクの異常度を計算し、その値が事前に定めた閾値を超えた場合に、異常値と判定する。

以上のようにして学習により推定したパラメータ θ から災害時に観測されたリンクの異常度を計算することが可能となり、道路ネットワークの交通状態を把握できる。

4. ETC2.0データを用いた交通状態の基礎分析

(1) 使用データの概要

本研究では、2015年9月1日から9月30日までの1ヶ月間のETC2.0プローブデータを用いて分析を行う。ETC2.0プローブデータはサーバ内での格納形態によって複数の様式に分類され、様式1は走行履歴情報や挙動履歴情報に関する生データが、様式2は様式1のデータをDRM区間単位で集計したデータが格納されている。本研究では1時間単位の平均リンク速度を分析対象とするため、様式2のうち、時間帯別平均旅行速度が格納されている様式2-3のデータを用いる。

(2) 分析対象地域

本研究では、2015年9月の関東東北豪雨で被災し、鬼怒川堤防決壊による浸水区域を含む茨城県南部から埼玉東部、千葉県北西部にかけての4つの2次メッシュの範囲を分析対象とする(図-1)。図-1では、9月11日時点での浸水範囲を青色で塗りつぶして表現してある。ただし、メッシュコード544000に一部含まれる高速道路(常磐自動車道)のデータに関しては、平常時と災害時でデータの挙動に大きな違いが見られなかったため、今回の分析対象からは除外した。

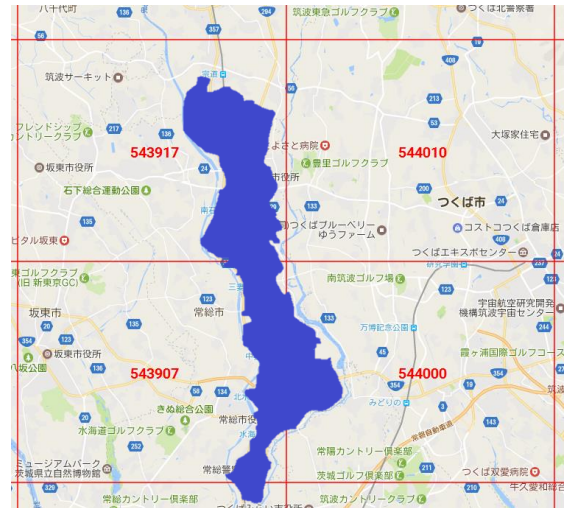


図-1 分析対象地域

(3) データセットの概要と前処理

本研究では、2015年9月の1ヶ月間のデータを災害発生前と災害発生後の期間で分類し、9月1日の6時台から9月8日の20時台の120データを前者、9月9日の6時台から9月30日の20時台の330データを後者とする。9月10日の正午頃に鬼怒川堤防が決壊しており、災害発生前の期間が少し短くなっているが、これは災害発生直前から災害発生直後にかけての変化を見るためである。データセットについては、分析対象地域で観測されたリンクのうち、災害発生前のデータ観測率が1割未満のリンクは今回の分析対象から除外した。作成した道路ネットワークの道路リンク数は1244本であり、平均リンク長は293mである。

(4) 欠損値に関する基礎分析

本研究のモデルの枠組みでは欠損値の異常を検知することはできないが、以下の手順で災害時に通行不可能であったリンクを推定することを試みる。リンクが遮断されると車両が通行できない状況が長時間継続し、連続的に欠損値が発生する。そのため、同一リンクにおいて災害時に平常時よりも継続的に欠損値が発生していれば、

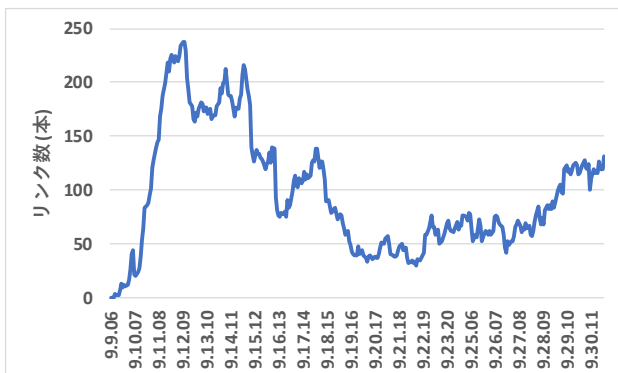


図-2 1時間単位の浸水判定リンク数の推移

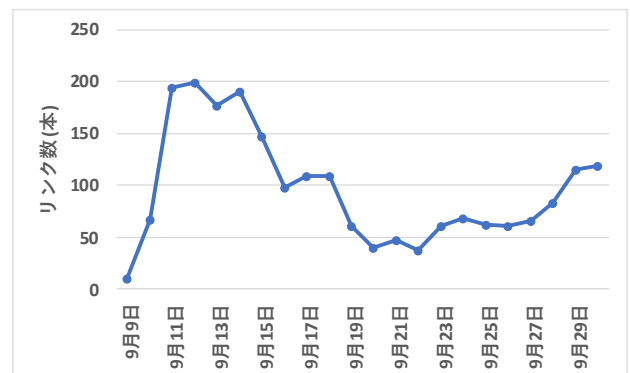


図-3 1日単位の浸水判定リンク数の推移

そのリンクは遮断している可能性があると考えられる。以上を踏まえて、各リンクについて災害発生前の最大連続欠損時間を算出し、災害発生後から各リンクについて改めて連続欠損時間を計算して、災害発生前の最大連続欠損時間を越えた段階で、その時間帯において該当リンクが遮断の可能性があるかと判定するモデルを考える。

図-2、図-3 は災害発生後のデータに遮断判定を実施した結果である。これらと図-1 を見比べると災害発生後に急激に浸水可能性リンク数が増加しており、浸水範囲が最大となった9月11日にかけて最大値をとっている。その後、浸水範囲の縮小とともに浸水可能性リンク数は減少に転じ、9月19日以降は落ち着いている。このように継続的な未観測リンクの発生の様子から、浸水範囲の広がりを一定程度把握することができると考えられる。

5. モデルの適用方法と適用結果

(1) モデルの適用方法

本研究では、災害発生前のデータを使って道路ネットワーク構造を学習し、学習されたモデルに災害発生後のデータを入力して異常度を計算することで、観測リンクの異常値を検出する。その際、リンクごとに異常度の閾値を設定する必要があるが、異常値の教師データがないため、そのままでは異常値判定を行うことができない。そこで、本研究では平常時に観測される可能性が低い速度を異常値と見なす。具体的には以下の方法で、リンクごとに異常度の閾値を定める。まず、災害発生前のデータを6つのグループに分割する。そのうち、5つを取り出した100の学習用データと残りの20の検証用データを作成する。学習用データにGLを適用して道路ネットワーク構造を学習し、学習されたモデルに検証用データを入力することで各リンクの異常度を全時間帯で計算する。この手続きを6回繰り返し、リンクごとに計算された異常度を昇順に並べ、95パーセンタイル値を異常度の閾値

とする。なお、欠損値については異常度の計算は行わず95パーセンタイル値の算定には含めていない。以上が、リンクごとに異常度の閾値を定める手順である。

その後、改めて災害発生前の全データを使い、道路ネットワーク構造を学習する。推定されたパラメータを使って、学習されたモデルに災害発生後のデータを逐次的に入力し、先ほど定めた異常度の閾値を超えたリンクを異常値と判定する。なお、GLを適用する際の正則化パラメータについては、0.5, 0.2 に設定して計算を行った。

(2) モデルの適用結果

a) 分析対象地域全体の傾向

図-4 は異常値判定されたリンクの本数を各時間帯で合計し、1日単位の平均値を時系列に並べた結果である。図-4の結果から、9月9日から災害発生後の9月12日にかけて異常リンク数が増加しているが、そこから9月14日にかけて減少し、9月17日にかけて再び増加に転じている。その後は緩やかな減少傾向となっているが、浸水範囲が大きく広がっていた災害発生直後に異常値が増加している傾向は見られない。災害発生直後は浸水範囲が広く、観測リンクの絶対数が少ないため、異常値判定されるリンク数が急激には増加しないと考えられる。そこで、各時間帯について異常リンク数を観測リンク数で割って異常判定率という指標を算出し、1日単位の平均値を時系列に並べた結果が図-5である。図-5の結果から、9月15日にかけて異常判定率が上昇し、浸水範囲の縮小に伴って異常判定率が減少している。しかしながら、災害発生直後の異常を早期に検出できているとは言い難い。

b) メッシュごとの傾向

水害によって影響を受けた範囲はメッシュによって異なるため、メッシュごとの差を検証する。図-6、図-7 は異常判定率をメッシュ別に集計した結果である。図-6の結果から、領域の面積に占める浸水範囲の割合が大きい543917と543907では災害発生前後で異常判定率に大

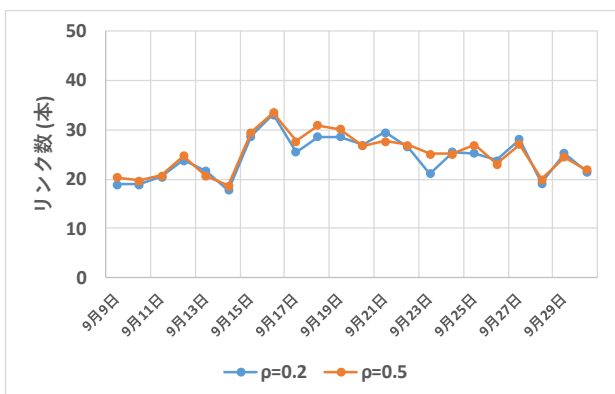


図-4 異常判定されたリンク数の推移

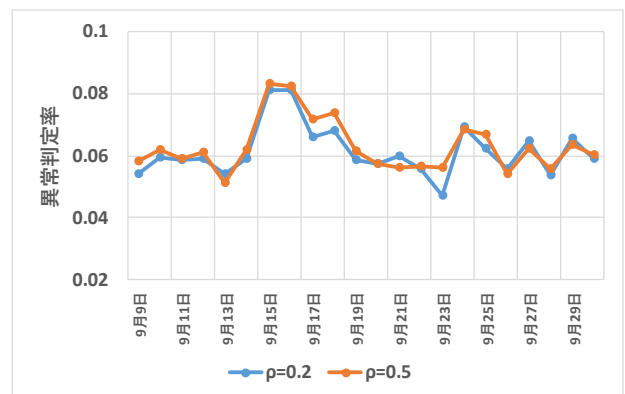


図-5 異常判定率の推移

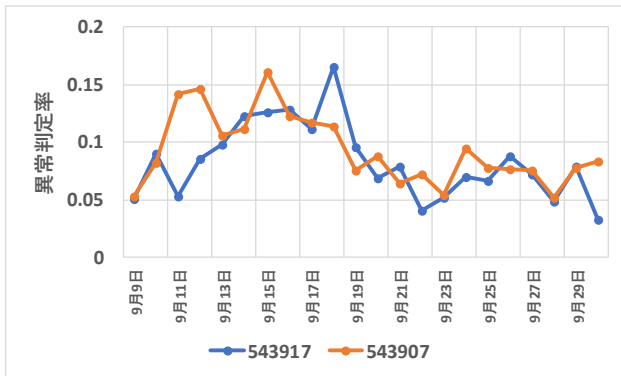


図-6 浸水範囲が広いメッシュの異常判定率の推移

大きな差があることが分かる。特に 543917 では、災害発生直後の 9 月 11 日、12 日にかけて異常判定率が急激に上昇しており、早期に一定数の異常値が検出されていると考えられる。一方、図-7 の結果から、領域の面積に占める浸水範囲の割合が小さい 544000 と 544010 においては、領域全体の傾向と同様に 9 月 15 日、16 日にかけて異常判定率が上昇しているが、前述の 2 つのメッシュと比較すると、災害発生直後で顕著な差は確認できなかった。

6. おわりに

本研究では災害時における道路ネットワークの相関構造の変化に着目し、観測リンクの異常値を検出するモデルを水害被災地域に適用した。被災地域全体で見ると、災害発生前後で異常値として検出されるリンク数が急激に増加することはなかったが、メッシュごとに見ると浸水範囲が広く分布していたエリアで観測データに占める異常値の割合が災害発生前後で急激に増加していた。以上の結果から、本モデルは災害時の交通状態把握に関して一定の精度を示していると考えられる。しかしながら、災害時に道路の通行可否を検知するためには本研究のモデルでは分析対象外の未観測リンクについても、異常の徴候を検出できることが望ましい。また、今回は外れ値検出を実装したが、相関構造の変化から確率変数間の依存性の変化を捉える構造変化検出を実装することで、より詳細な交通状態把握が可能になると考えられる。以上の方針で研究を発展させることが今後の課題である。

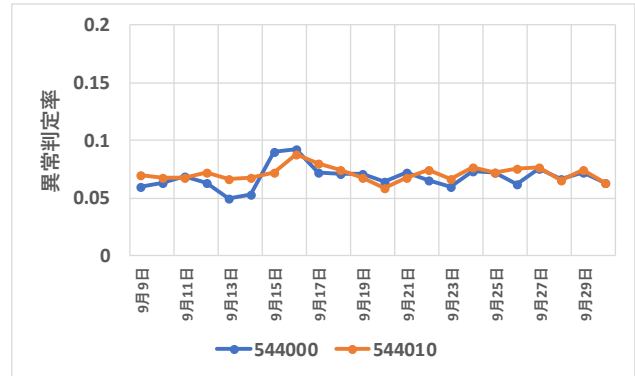


図-7 浸水範囲が狭いメッシュの異常判定率の推移

参考文献

- 1) 原祐輔, 花岡洋平, 桑原雅夫: 道路道路ネットワーク内の関係性に着目した長期観測プローブデータによるプローブ未観測リンクの交通状態補間, 交通工学論文集, 第2巻, 第1号, p.1-10, 2016
- 2) ETC の利用状況, 国土交通省, <http://www.mlit.go.jp/road/yuryo/etc/riyou/index.html>
- 3) 秦康範, 鈴木猛康, 下羅弘樹, 目黒公郎, 小玉乃理子: 新潟県中越沖地震における通れた道路マップの提供とプローブカー情報の減災利用実現に向けた課題と展望, 日本地震工学論文集, 第9巻, 第2号 (特集号) 2009
- 4) 水谷友彰, 伊坪慎二, 鳥海大輔, 牧野浩志: ETC プローブデータ活用による災害時通行可能ルートの把握可能性検討, 第14回 ITS シンポジウム 2016
- 5) 丸橋弘治, 山影譲: テンソル分解を用いた交通状態推定による異常状態検知, DEIM Forum 2016, H8-3
- 6) 増田淳基, 松平正樹: ETC2.0 プローブデータを利用した交通異常検知, 第33回人工知能学会全国大会論文集, 2019
- 7) 稲場亘, 柳沼秀樹, 寺部慎太郎, 康楠, 田中皓介: ETC2.0 プローブデータを用いた災害時の交通道路ネットワーク状態推定, 土木計画学研究・講演集, 2019
- 8) 上野玄太: データ同化における大規模グラフィカルモデルの推定について, 統計数理, Vol.61, No.1, pp.17-46, 2013
- 9) Friedman, J., Hastie, T. and Tibshirani, R.: Sparse inverse covariance estimation with the graphical lasso, Biostatistics, Vol.6, pp.2125-2149, 2012
- 10) Mazumder, R., and Hastie, T.: The graphical lasso: New insights and alternatives, Electronic journal of statistics, Vol.6, pp.2125-2149, 2012
- 11) 井手剛, 杉山将: 異常検知と変化検知, 講談社, 2016.

(Received October 2, 2020)

(Accepted October 2, 2020)

ROAD NETWORK ANOMALY DETECTION AT THE HEAVY RAIN DISASTER USING ETC2.0 PROBE DATA

Wataru INABA, Hideki YAGINUMA, Shintaro TERABE and Kosuke TANAKA