

LTE 環境における応答遅延特性の時系列モデリングによる分析

山本 航平[†] 若宮 直紀[†] 中野 亮^{††} 藤原 亮介^{††}

[†] 大阪大学大学院情報科学研究科バイオ情報工学講座

〒565-0871 大阪府吹田市山田丘 1-5

^{††} 株式会社日立製作所 研究開発グループ

〒185-8601 東京都国分寺市東恋ヶ窪 1-280

E-mail: [†]{k-yamamt,wakamiya}@ist.osaka-u.ac.jp, ^{††}{ryo.nakano.xd,ryosuke.fujiwara.mb}@hitachi.com

あらまし 産業用モニタリングシステムにおける運用管理コストの低減のため、迅速な障害検知・予測、ならびに原因の特定と対処法の提示が求められている。その実現にむけて、無線機器からサーバまでの LTE 回線を含む通信路について、異なる時間帯において応答遅延の計測を行った。本稿では、遅延の変動特性や、曜日や時間帯に依存した傾向の有無について、時系列モデルを用いたクラスタリングによって分析した結果を示す。

キーワード Long Term Evolution, 応答遅延, 時系列モデリング, 異常検知

Analysis of delay characteristics of connection through LTE by time series analysis

Kohei YAMAMOTO[†], Naoki WAKAMIYA[†], Ryo NAKANO^{††}, and Ryosuke FUJIWARA^{††}

[†] Department of Bioinformatic engineering, Graduate school of Information Science and Technology, Osaka University

1-5 Yamadaoka, suite-shi, Osaka, 565-0871, Japan

^{††} Research & Development Group, Hitachi, Ltd.

1-280 Higashikoigakubo, kokubunji-shi, Tokyo, 185-8601, Japan

E-mail: [†]{k-yamamt,wakamiya}@ist.osaka-u.ac.jp, ^{††}{ryo.nakano.xd,ryosuke.fujiwara.mb}@hitachi.com

Abstract In order to reduce the costs of operation and management in industrial monitoring systems, it is necessary to detect or predict failures quickly, identify the cause, and present countermeasures. To achieve this, we measured response delays of connection from one wireless device to one server, including LTE, in different time zones. In this technical report, we show the results of analyzing the fluctuation characteristics of response delay and the tendency depending on the day of week or time zone by clustering using a time series analysis.

Key words Long Term Evolution, response delay, time series analysis, anomaly detection

1. はじめに

近年, IoT (Internet of Things) 技術の発展とともに産業用モニタリングシステム [1] が普及している。これは、工場などの産業施設に設置された機器から直接、あるいは配置したカメラやセンサーなどの IoT デバイスを通じて、機器の稼働状態に関する情報をリアルタイムで収集し、キャリア回線を通じてクラウドサーバに送信、さらにクラウドサーバでデータ処理を行い、運用管理担当者に提示するものである。従来作業員が巡回し行っていた工場内の機器の点検業務を自動化できるため、人員コストの軽減、目視確認より生じる人的ミスの低減、リア

ルタイムなデータの利活用などの効果が期待できる。一方で、長期間の運用のなかで機器に具備された IoT デバイスの故障、工場内ネットワークの通信の途絶、クラウドサーバへの通信の遅延などの障害は避けられない。このような障害が発生した場合には工場内の機器の稼働状況を把握することが困難となり、工場の稼働停止や業務の遅れなどが引き起こされ大きな損失をもたらす可能性がある。

障害発生時には迅速な復旧作業が求められるが、障害はその原因や内容、規模によってさまざまである。しかしながら、現状ではシステムから得られる情報を用いてこれらの障害を適切かつ迅速に区別することが困難であるため、障害発生時には運用

管理担当者が直接現場に行き障害の原因や内容、規模を確認する必要がある、多大な運用管理コストが発生している。そのため、モニタリングシステムによる運用管理コストの低減のためには、直接また間接的に取得できる情報にもとづいて、障害発生を迅速に検知また予測するとともに、障害の原因を特定し、加えてその対処法を示すことが求められている。

すでに我々の研究グループでは機器間で測定される受信電波強度の時間変化にもとづいて空間的な電波伝搬特性の変動を推定することにより工場内での無線通信の異常を検知する手法を提案している [2]。そこで、本報告では工場内で機器稼働情報を収集する無線機器からクラウドサーバまでの通信路で発生する異常の検知にむけてさまざまな曜日、時間帯で応答遅延を計測し特性の分析を行う。具体的には、キャリア回線として産業用モニタリングシステムに広く用いられている LTE (Long Term Evolution) 回線を用いる。LTE 回線に関する研究として、応答遅延時間を他の無線回線と比べて評価を行った研究 [3] [4] や LTE 環境下での TCP パケットの振る舞いに関する調査 [5]、LTE 環境の応答遅延において、人の混雑状況にかかわらず発生する低遅延帯の分布が部分正規分布に従うことを示唆した研究 [6] などが行われている。しかし、LTE を利用したインターネットとの通信において発生する遅延などの異常を検知・予測するといった観点からの通信特性の多角的な分析は十分になされていない。そこでまずは分析に用いる実測データの収集として、月曜日から日曜日の 3 時、7 時、12 時、17 時、20 時のそれぞれ 1 時間において私が在籍している研究室に設置した無線端末から AWS サーバへ ping による応答遅延の計測を行った。計測した応答遅延に、例えば曜日、時間帯に依存した傾向が認められれば、その傾向と対比することによって応答遅延の通常ではない振る舞い、すなわち異常を検知することができる。そこで、時系列モデリングによる回帰で得られたパラメータやその主成分を用いてクラスタリングすることで計測データをグループ化し、共通する特徴を見いだす。

第 2 章では本報告で実施した計測実験の設定について述べる。第 3 章では時系列モデルによる回帰について述べる。第 4 章では回帰結果に用いたクラスタリングにもとづく分析について述べる。第 5 章では全体のまとめと今後の課題について述べる。

2. 計測実験の設定

本報告では図 1 に示す構成で計測実験を行った。モニタリングシステムにおける無線端末としては LTE モジュールとして Quectel 社製 EC21-J を搭載した Raspberry Pi を用いた。また、LTE 回線としては IIJ モバイル社のサービスタイプ D 定額プランライト (いちねん プリペイド) を用いた。IIJ モバイル社は他の通信事業者から通信回線を借り受け、サービスを提供している MVNO (Mobile Virtual Network Operator) であり、サービスタイプ D では NTT ドコモ社の回線を使用している。月あたり通信量が 3GB を超過すると通信速度が 256kbps に制限されるが、本実験中には速度制限は課されなかった。

クラウドサーバとしては本計測を行うにあたり契約した一台

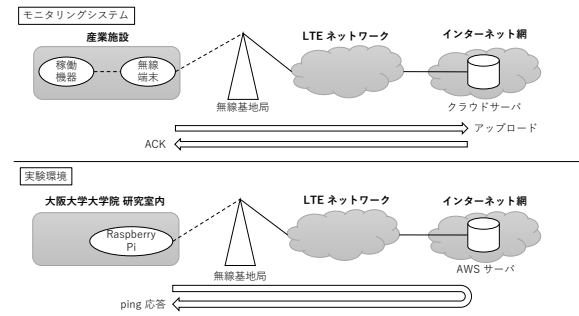


図 1 モニタリングシステムと実験環境の対応図

の AWS サーバを用い、大阪大学敷地内の研究室に設置した Raspberry Pi から ping を用いて応答遅延を計測した。自動的に計測データを取得できるように、Raspberry Pi 上で動作する Raspbian において計測用のスクリプトを用いることで、15 秒毎に時刻を取得し、続けて ping (パケットサイズ 60 バイト、ICMP ECHO メッセージ、パケット数 1) を実行した。計測時刻、ping の出力をログデータとして取り出し、分析を行った。また、日時が応答遅延に与える影響を調べるため、様々な時間帯で計測を行った。計測は 2/29 (土) から 3/27 (金) までの四週間に渡って行った。

3. 時系列モデルによる回帰

本報告では、時系列モデルとして式 (1) と式 (2) で表される ARMA-GARCH (Autoregressive Moving Average - Generalized Autoregressive Conditional Heteroscedasticity) モデル [7] を用いる。

$$y_t = \sum_{i=1}^p a_i y_{t-i} + \sum_{i=1}^q b_i \varepsilon_{t-i} + c + \varepsilon_t \quad (1)$$

$$\varepsilon_t \sim N(0, h_t) \quad \text{i.i.d}$$

$$h_t = \omega + \sum_{i=1}^r \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^s \beta_i h_{t-i} \quad (2)$$

各一時間の計測実験において Raspberry Pi が AWS サーバから受け取る ping の応答遅延の実測値を計測時刻順に y_1, y_2, \dots, y_N と表す。式 (1) において、 c は定数項、 ε_t はノイズ項であり平均 0 分散 h_t の正規分布に従いそれらは互いに独立であるとする。時刻 t における実測値 y_t は、過去の p 時点前までの実測値の線形和と q 時点前までのノイズ項の加重和と自身のノイズ項によって表される。また、式 (2) において、 ω は定数項であり、時刻 t におけるノイズ項が従う正規分布の分散 h_t は、過去の r 時点前までのノイズ項の線形和と s 時点前までのノイズ項が従う正規分布の分散の加重和によって表される。

各パラメータのパラメータ数 (p, q, r, s) は次数と呼ばれ、精度の良い回帰を行うためには対象とする時系列データに応じてこの次数を適切に定める必要がある。しかしながら、計測データのクラスタリングによる分類のためには異なる曜日や時間帯での計測によって得られた時系列データに対して次数を統一し、

て回帰しなければならない。そこで、本報告では、まず、各計測データに対し ARMA-GARCH モデルの次数を変えながら適用し、AIC（赤池情報量規準）[8][9] によってその妥当性を評価することで、各時系列データにとって適切な次数を求める。ここで、AIC は式 (3) で定められる指標である。高すぎる次元での回帰による過適合の問題を考慮し、時系列モデルによる回帰のよさを評価する第 1 項とパラメータの少なさを評価する第 2 項を組み合わせている。AIC が小さいほどよいモデルとされている。

$$AIC = -2 * (\text{対数尤度}) + 2 * (\text{パラメータ数}) \quad (3)$$

また、対数尤度は式 (4) と式 (5) で定められる。

$$(\text{対数尤度}) = \sum_{t=1}^N \log \left[\frac{1}{\sqrt{2\pi h_t}} \exp \left\{ -\frac{1}{2h_t} (y_t - \hat{y}_t)^2 \right\} \right] \quad (4)$$

$$\hat{y}_t = \sum_{i=1}^p a_i y_{t-i} + \sum_{i=1}^q b_i \varepsilon_{t-i} + \varepsilon_t \quad (5)$$

次に、各時系列データ間で AIC による最適次数を比較し、その最大のものを用いることとする。次数の探索範囲は $0 \leq p \leq 2$, $0 \leq q \leq 2$, $r=1$, $0 \leq s \leq 1$ とした。

各時系列データをその最適次数ごとにまとめると表 1 となっていたため、次数を $(p, q, r, s) = (2, 2, 1, 1)$ で統一した。この次数での ARMA-GARCH モデルによる回帰結果の一部を 2 に示す。

また、応答遅延の変動に注目し、応答遅延の実測値の差分系列 $\{\Delta y_t | y_t = y_{t-1}\}$ に対しても同様の手法でモデルの適用を行った。この各変動値の時系列データをその最適次数ごとにまとめると表 2 となっていたため、次数は $(p, q, r, s) = (2, 2, 1, 1)$ で統一を行った。この次数での変動値に対する ARMA-GARCH モデルでの回帰結果の一部を図 3 に示す。

また、最大次数で統一することにより本来の最適次数よりも大きな次数で回帰が行われる時系列データがあるが、次数を最適なものに比べ大きく設定することがモデルの回帰精度に際立った悪影響を与えないことを対数尤度を比較することにより確認した。一例として、表 3 に最適次数と統一次数が異なるある一つの時系列データに対して、それぞれの次数で回帰を行った時のモデルの対数尤度を示す。

図 2 より、赤線で示される ARMA-GARCH モデルによる回帰線は青線で示された ping 応答の実測値の推移の傾向を大まかに捉えられている。したがって、応答遅延時間が時間とともに単調に増加するなどの異常は、例えば工場内部の無線端末上でリアルタイムに ARMA-GARCH モデルによる応答遅延データの回帰を行うことで断続的に定まるパラメータの変化や、現場環境における通常時の傾向を反映した ARMA-GARCH のパラメータを用いて行った予測値から実測値が大きく外れ続けた場合などをもとに、検知することが可能と考える。しかしながら一方で、単発的に発生する大きな応答遅延を予測できていない。これは、この単発的に発生する応答遅延に周期性や過去データからみられる兆候がなく、予測が困難であることを示唆

表 1 最適次数と対応する実測値データの数

次数 (p, q, r, s)	データ数
(0,0,1,0)	9
(0,0,1,1)	6
(0,1,1,0)	11
(0,1,1,1)	4
(0,2,1,0)	5
(0,2,1,1)	3
(1,0,1,0)	3
(1,0,1,1)	3
(1,1,1,0)	4
(1,1,1,1)	1
(1,2,1,0)	6
(1,2,1,1)	8
(2,0,1,0)	4
(2,0,1,1)	3
(2,1,1,0)	3
(2,1,1,1)	1
(2,2,1,0)	21
(2,2,1,1)	11

表 2 最適次数と対応する変動値データの数

次数 (p, q, r, s)	データ数
(0,0,1,0)	0
(0,0,1,1)	0
(0,1,1,0)	11
(0,1,1,1)	4
(0,2,1,0)	7
(0,2,1,1)	3
(1,0,1,0)	0
(1,0,1,1)	0
(1,1,1,0)	4
(1,1,1,1)	2
(1,2,1,0)	21
(1,2,1,1)	10
(2,0,1,0)	0
(2,0,1,1)	0
(2,1,1,0)	2
(2,1,1,1)	0
(2,2,1,0)	19
(2,2,1,1)	23

3/6 (金) 17:00-18:00, 3/7 (土) 17:00-18:00
3/12 (木) 12:00-13:00, 3/12 (木) 17:00-18:00
3/13 (金) 17:00-18:00, 3/14 (土) 17:00-18:00
3/15 (日) 17:00-18:00, 3/16 (月) 17:00-18:00
3/17 (火) 20:00-21:00, 3/20 (木) 17:00-18:00
3/22 (土) 17:00-18:00, 3/23 (日) 12:00-13:00
の計測データは Raspberry Pi の動作不良のため未使用

表 3 最適次数と統一次数でのモデルの尤度

	最適次数での対数尤度	統一次数での対数尤度
実測値データ	-969.8327	-971.9196
変動値データ	-924.6495	-922.7543

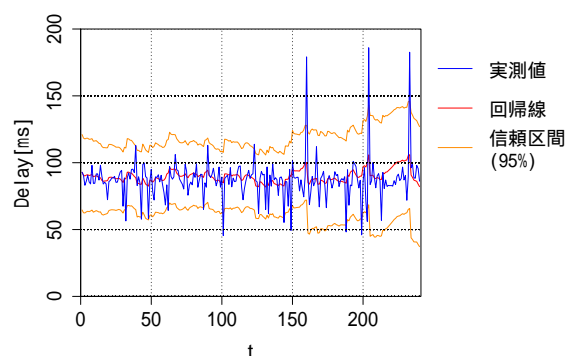


図 2 3/1 (日) 12:00-13:00 計測の実測値に対する次数 (2,2,1,1) の ARMA-GARCH モデルによる回帰

している。この単発的な応答遅延によって、モデルの回帰精度が悪化したり、その発生頻度などに関する異常を検知できなかったりすることが考えられ、実用的な異常検知にむけてはこれらの点を改善するような別の手法と組み合わせることが必要だと思われる。また、図 3 においては図 2 の場合と同様に、回帰線が変動値の推移をおおよそ捉えられ単発的な変動は捉えら

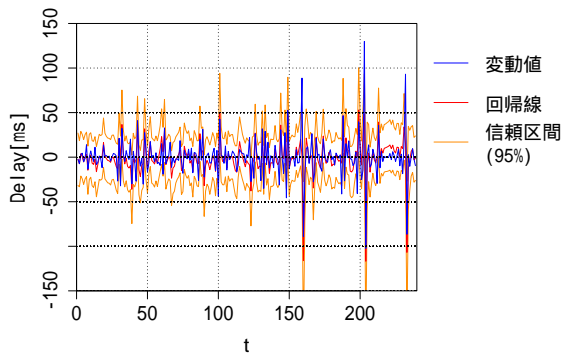


図 3 3/1 (日) 12:00-13:00 計測の変動値に対する次数 (2,2,1,1) の ARMA-GARCH モデルによる回帰

れていなかったものの、大きな ping 応答に対する立下りは捉えられていた。このことから、変動値に対する ARMA-GARCH モデルではこれらの大きな応答遅延の単発性を表現できていると考えられる。従って、大きな応答遅延が連続して発生するような異常は ARMA-GARCH モデルを用いることで検知可能ではないかと考えられる。

4. クラスタリングによる分類

ARMA-GARCH モデルを実測値データもしくは変動値データに対して適用した結果定まるパラメータ W をもとにクラスタリングを行い、時間帯や場所ごとの ping 応答遅延特性を分析する。また、 W は、実測値データと差分系列データともに

$$W = [c, a_1, a_2, b_1, b_2, \omega, \alpha_1, \beta_1]$$

で与えられる。

文 献

- [1] A.C. Lima-Filho, R.D. Gomes, M.O. Adissi, T.A.B. Da Silva, F.A. Belo, and M.A. Spohn, "Embedded system integrated into a wireless sensor network for online dynamic torque and efficiency monitoring in induction motors," IEEE ASME transactions on mechatronics, vol.17, no.3, pp.404–414, 2012.
- [2] 辰巳 陸, "実環境における Radio Tomographic Imaging によるイベント検知の検証," 特別研究報告, 大阪大学, Feb 2019.
- [3] A. Striegel, S. Liu, X. Hu, and L. Meng, "Lte and wifi: Experiences with quality and consumption.," FNC/MobiSPC, pp.418–425, 2014.
- [4] J. Garcia, S. Alfredsson, and A. Brunstrom, "Delay metrics and delay characteristics: A study of four swedish hsdpa+ and lte networks," 2015 European Conference on Networks and Communications (EuCNC)IEEE, pp.234–238 2015.
- [5] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z.M. Mao, S. Sen, and O. Spatscheck, "An in-depth study of lte: effect of network protocol and application behavior on performance," ACM SIGCOMM Computer Communication Review, vol.43, no.4, pp.363–374, 2013.
- [6] 田行里衣, 池上 大介, "LTE 環境下の実測データに基づく RTT 分布モデリング," 電子情報通信学会, vol.118, no.140, pp.57–61, 2018.
- [7] C.G. Lamoureux and W.D. Lastrapes, "Persistence in variance, structural change, and the garch model," Journal of Business & Economic Statistics, vol.8, no.2, pp.225–234, 1990.
- [8] H. Bozdogan, "Model selection and akaike's information cri-

terion (aic): The general theory and its analytical extensions," Psychometrika, vol.52, no.3, pp.345–370, 1987.

- [9] K.P. Burnham and D.R. Anderson, "Multimodel inference: understanding aic and bic in model selection," Sociological methods & research, vol.33, no.2, pp.261–304, 2004.