管網端部での水圧低下に基づく漏水発生予測に向けた機械学習の利用

Machine learning for predicting leakage occurrence based on water pressure observation

22WM1315 安江　崇志　　指導教員 丸山　喜久
Takashi Yasue

SYNOPSIS

Currently, more than 20,000 water leakage and breakage incidents occur annually in Japan's water supply system. Leaks in water pipes can be roughly classified into two types: aboveground leaks that flow out above the ground and underground leaks that do not flow out above the ground but flow underground. While aboveground leaks are easy to detect because they are visible, underground leaks cannot be directly confirmed visually. Therefore, development for early detection method is required. In this study, the authors tried to detect leakage location based on water pressure observation, assuming the monitoring of pipelines using smart meters, which are currently in widespread use.

1. はじめに

我が国の上水道では年間 2 万件を超える漏水・破損事故が発生している．上水道管路の法定耐用年数は約 40 年とされており，特に高度経済成長期に整備された施設の老朽化に伴う漏水事故が懸念されている．

水道管の漏水は，地上に流れ出す地上漏水と，地上には流れ出さず地下で流れている地下漏水の 2 種類に大別できる．東京都水道局によると，令和 2 年度は地上漏水に伴う修理対応が 6,790 件，地下漏水に伴う対応が 306 件であった [1]．地上漏水は人目に触れるため発見しやすいが，地下漏水は漏水の状況を直接目視で確認できず長時間地中に漏水し続けることがある．地下漏水の発見には，人が漏水の音を聴き取る音聴法や夜間流量測定法などがあるが，熟練の技術を要することや，計画作業となるため時間がかかることが課題となっている．地下漏水の発見が遅れると，道路陥没等の大きな事故につながることが懸念され，地下漏水の早期発見のための技術開発が求められている．

そこで本研究では，現在普及が進んでいるスマートメータによる管路モニタリングを想定する．スマートメータに圧力センサが搭載されたものと仮定し，管網端部の多点における水圧の計測データに基づき，上水道管路の漏水を予測する機械学習モデルの構築を目的とする．スマートメータとは，通信機能を備えた水道メータのことを指す．スマートメータが導入されると，高頻度に検針データを得ることができる．スマートメータには様々なセンサを搭載することが検討されており，圧力センサもその 1 つである．本研究では，各家庭に備え付けられたスマートメータから得られる水圧を活用し，漏水位置の推定を目指す．

まず，多数の漏水シナリオを想定した管網解析を実行し，スマートメータ設置位置での水圧を算出する．これらを学習データとして機械学習し，管網内の漏水有無を検知する漏水検知モデル，漏水箇所を予測する漏水予測モデルを構築する．

2. 管網解析

(1) 管網解析の概要

本研究では，埋設管網を対象とした管網解析 [2]を行い，管網内の 1 箇所で漏水が発生した場合の管網端部での水圧変化を計算する．管網解析とは，配水池の水位，配水管の敷設状況，口径，節点流出水量等の与条件下で，配水管の中をどのように水道水が流れるかシミュレーションし，管路流量，節点水圧を求めることをいう．

各管路の損失水頭 H を流量 Q で表すための損失水頭式（ヘーゼン・ウィリアムスの式）[3]は，以下の式(1a), (1b)のように書ける．

$$H = r \cdot Q^{1.85} \tag{1a}$$

$$r = 10.666 \cdot C^{-1.85} \cdot D^{-4.87} \cdot L \tag{1b}$$

ここで H は損失水頭(m)，C は流速係数，L は管路長(m)，Q は流量(m^3/s)，D は口径(m)である．

各節点に接続する管路の Q が満たすべき節点方程式は式(2)のように書ける．

$$\sum \pm Q_i = p_i \tag{2}$$

ここで i は配水基地以外の節点，p は節点 i からの流出水量である．

各要素閉管路を構成する管路の H が満足すべき閉管路方程式は，式(3)のように書ける．

$$\sum \pm H = 0 \tag{3}$$

ここで H は損失水頭である．

(2) 管網図の作成

本研究では，実際の埋設管網を参考に管網解析を行う．対象地域内の全ノード数は 424 であり，そのうち配水管と給水管の分岐点のノード数は 274 である．本研究では各家庭のスマートメータ設置位置を，配水管と給水管の分岐点に置き換えて扱う．また，対象地域内には 6 種類の口径の配水管が存在する．配水管と給水管の分岐点，配水管の口径の分布を**図-1** に示す．

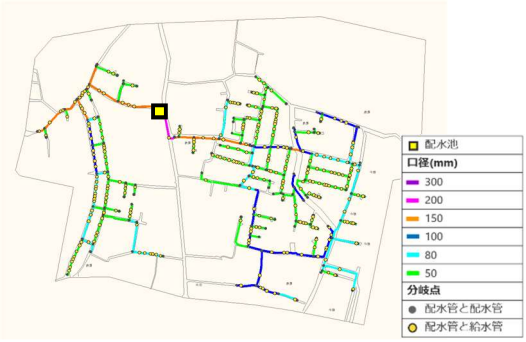

図-1 対象地域の配水網内の口径と分岐点の分布

(3) データセットの作成

本研究では，以下の手順に従ってデータセットを作成した．

a) 通常時（漏水なし）のシナリオデータの作成

配水管と給水管の分岐点のノードに対し 0～0.05 m^3/h の取り出し水量をランダムに設定する．この水量は，概ね洗濯の使用水量に相当する [4]．地域内に配水池を 1 箇所設定し，60 m の水頭固定水源とした．

b) 漏水時のシナリオデータの作成

本研究では，管網内に 0.01 m^3/h～10 m^3/h の漏水が 1 箇所で発生したことを想定する．まず，漏水点を，管網内の 424 地点のノードの中から 1 箇所選定する．次にそのノードに対し 0.01 m^3/h～10 m^3/h の範囲でランダムに漏水量を設定する．このように漏水を設定した状態で管網解析を行い，各メータでの通常時との水圧変化量を求める．

3. 機械学習に基づく漏水検知モデルの構築

(1) 使用するデータセット

本研究では実際の漏水状況をもとに，毎日夜間に 1 回計測することを想定しデータセットを作成した．本研究で仮定した鋼管，鋳鉄管の漏水率を**表-1** に示す．**表-1** は，千葉県営水道の給水エリア内で 2013 年から 2021 年までの 8 年間に発生した配水管の漏水件数，管種ごとの敷設延長をもとに算出したものである．これらは，いずれもヒアリング調査等に基づき設定した．

本研究では，**表-1** から 1 日ごとの漏水率を算出し，その漏水率を 100 倍にした値を基に通常時のシナリオデータ，漏水時のシナリオデータの割合を設定した．ここで漏水率を 100 倍としたのは，給水管での漏水は配水管の数十倍多いというヒアリング調査で得られた実績に基づくものである．本研究では，通常時の管網解析を 9980 シナリオ，漏水時の管網解析を 20 シナリオ実行し，それらを合わせた 10000 シナリオをデータセットとして用いた．

表-1 管種ごとの漏水実績

	鋳鉄管	鋼管	その他	合計
漏水件数(件)	50	13	19	82
敷設距離(km)	8743	167	296.7	9207.2
漏水率 (件/km)	0.0057	0.0777	0.06404	0.0089

(2) 機械学習手法

本研究での漏水検知では，機械学習手法に異常検知でよく用いられている Local Outlier Factor[5]（LOF）を用いる．

(3) 精度評価指標

本研究では予測結果を評価する上で，正答率，適合率，再現率，F 値の 4 つの指標を導入する．各指標式は式(4)～(7)に示す．式中の TP, FP, FN TNは，**表-2** の混同行列のように定める．

正答率 Acc は，全データに対して以下のように算出される．

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \tag{4}$$

適合率 Pre は，予測モデルが漏水ありと予測したデータのうち，実際に漏水ありのときの割合を示す．誤検知の割合を調べるときに使用する．

$$Pre = \frac{TP}{TP + FP} \tag{5}$$

再現率 Rec は，実際に漏水ありのデータのうち，予測モデルが漏水ありと正しく予測した割合を表す．漏水を見逃す割合を調べるときに使用する．

$$Rec = \frac{TP}{TP + FN} \tag{6}$$

F 値 F は，適合率と再現率の調和平均である．これらの 2 指標を総合的に見るときに使用する．

$$F = \frac{2 \cdot Pre \cdot Rec}{Pre + Rec} \tag{7}$$

表-2 精度評価のための混同行列

		予測結果	
		漏水なし	漏水あり
実際の結果	漏水なし	TN	FP
	漏水あり	FN	TP

(4) 学習条件と予測モデルの構築

本研究では，管網端部の水圧変化から配水網内において漏水が発生しているか判別するモデルを構築した．本研究の漏水検知モデルでは，各ノードの水圧変化率，水圧変化量を説明変数とし，漏水発生の有無を予測する．

スマートメータは現在普及が進んでいる段階であり，これから段階的に導入される．そこで，地域内のスマートメータ導入率が低い場合を想定し，メータの導入率を 50%，25%，10%の場合を仮定し，漏水予測モデルの精度評価を行った．具体的には，メータの導入率を 100%， 50%，25%，10%と変化させ，ランダムにメータ設置点を設定した．それぞれモデル A，モデル B，モデル C，モデル D とした．機械学習手法には LOF を採用する．LOF では，近傍の点の個数である k の値がハイパーパラメータであり，人間が設定する必要がある．しかし，k の値を大きくすると「データの周辺のデータの疎密の度合い」から乖離し異常値を検出できなくなることがある．一方で，k の値が小さすぎると外れ値による影響を強く受けることがある．本研究では，ハイパーパラメータ k の値を 0～300 の間でグリッドサーチを行い，モデルごとに適切な k の値を設定した．

(5) 漏水検知モデルの評価

表-3，**表-4** にモデル A～D のデータセットに対する精度を示す．

表-3 漏水検知モデルの結果の混同行列

モデル A（導入率 100%）		予測結果	
		漏水なし	漏水あり
実際の結果	漏水なし	9953	27
	漏水あり	2	18

モデル B（導入率 50%）		予測結果	
		漏水なし	漏水あり
実際の結果	漏水なし	9953	27
	漏水あり	2	18

モデル C（導入率 25%）		予測結果	
		漏水なし	漏水あり
実際の結果	漏水なし	9928	52
	漏水あり	2	18

モデル D（導入率 10%）			
		漏水なし	漏水あり
実際の結果	漏水なし	9451	529
	漏水あり	2	18

表-4　漏水検知モデルの結果の評価指標

	モデル A（導入率 100%）	モデル B（導入率 50%）	モデル C（導入率 25%）	モデル D（導入率 10%）
正答率 Acc	0.997	0.997	0.995	0.947
適合率 Pre	0.400	0.400	0.257	0.033
再現率 Rec	0.900	0.900	0.900	0.900
F 値 F	0.554	0.554	0.400	0.063

どのモデルにおいても正答率は 0.9 を超えるような高い値となった．これは，混同行列における TN（漏水なしと予測したもののうち実際に漏水なしであったもの）が非常に多いことから正答率が高くなったと考えられる．本研究では漏水検知を目的としていることから，再現率に着目すると，どのモデルでも再現率 0.9 となり，概ね漏水は検知できているといえる．漏水検知ができなかった FN（漏水なしと予測したもののうち，実際には漏水ありであったもの）のシナリオを調べると，漏水量が 0.2 m^3/h 未満であった．漏水量が小さい場合には検出が困難であるものの，0.2 m^3/h 以上の漏水を検知する場合には，メータ設置数が少ない場合でも検知できる可能性が示された．しかし，スマートメータの導入率が 10%のときは，FP（実際には漏水がないにもかかわらず，漏水が発生していると誤って検知したケース）が 529 あり，F 値も 0.063 と非常に小さい．導入率が 25%，50%となると誤検知のケース数（FP）が 52，27 と改善されており，スマートメータの観測値に基づく漏水検知を実現させるには，導入率がある程度高い必要がある．

4.　機械学習に基づく漏水予測モデルの構築

(1) 使用するデータセット

本研究では 2 種類のデータセットを作成した．データセット A では漏水点をシナリオごとにランダムに設定した．データセット B では，管網内の上水道管の管種に応じて漏水箇所を設定した．具体的には，鋼管，鋳鉄管の漏水率の 1000 倍を平均値としたポアソン乱数をシナリオごとに発生させ，確率変数の最も大きなノードを漏水点に設定した．本研究で仮定した鋼管，鋳鉄管の漏水率は表-1 をもとにそれぞれ設定した．

(2) 機械学習手法

本研究では，前章の管網解析で得られたシミュレーション結果に基づき機械学習を行い，管網内の漏水箇所を予測する漏水予測モデルを構築する．機械学習手法には，Random Forest[6]，XGBoost[7]，LightGBM[8]を用いる．

(3) 精度評価指標

本研究では予測結果を評価する上で，平均正答率，平均適合率，平均再現率，平均 F 値の 4 つの指標を導入する．これらの評価指標は**表-2** に示す混同行列の$TP, FP, FN\ TN$をメッシュごとに算出し，それらを式(8)～(11)にしたがって算出したものである．

$$Acc = \frac{\sum TP_i + \sum TN_i}{\sum TP_i + \sum FP_i + \sum FN_i + \sum TN_i} \tag{8}$$

$$Pre = \frac{\sum TP_i}{\sum TP_i + \sum FP_i} \tag{9}$$

$$Rec = \frac{\sum TP_i}{\sum TP_i + \sum FN_i} \tag{10}$$

$$F = \frac{2 \cdot Pre \cdot Rec}{Pre + Rec} \tag{11}$$

(4) 学習条件と予測モデルの構築

本研究では，配水網を 50m メッシュで区切り，管網端部の水圧変化から各メッシュにおいて漏水点を含むかどうかを判別するモデルを構築した．2 章で述べたデータセット A，B と 3 章で述べた 3 種類の機械学習手法を組み合わせた 6 つのモデルの構築を行った．本研究の漏水予測モデルでは，各ノードの水圧変化率，水圧変化量，管種（モデル B のみ）を説明変数とし，漏水点を含むメッシュを予測する．

(5) 漏水予測モデルの評価

a) データセット A：漏水点をランダムに設定した場合の予測精度

表-5 にモデル A1～A3 のテストデータに対する精度を示す．

表-5 モデル別の予測精度

	モデル A1	モデル A2	モデル A3	モデル B1	モデル B2	モデル B3
機械学習手法	Random Forest	XGBoost	LightGBM	Random Forest	XGBoost	LightGBM
平均正答率	0.993	0.993	0.994	0.999	0.999	0.999
平均適合率	0.81	0.816	0.832	0.958	0.935	0.962
平均再現率	0.403	0.388	0.538	0.94	0.927	0.955
平均 F 値	0.538	0.526	0.654	0.949	0.932	0.959

どのモデルにおいても平均正答率は 0.99 を超えるような非常に高い値となった．これは，混同行列における TN_i（漏水なしと予測したもののうち実際に漏水なしであったもの）が非常に多いことから平均正答率が高くなったと考えられる．本研究では漏水検知を目的としていることから，平均再現率に着目する．それぞれのモデルの平均再現率を比較すると，モデル A1 が 0.403，モデル A2 が 0.388，モデル A3 が 0.538 となり，機械学習手法に LightGBM を用いたモデル A3 が最も良い結果を示した．

b) データセット B：管種に応じて漏水シナリオを作成した場合の予測精度

表-5 にモデル B1～B3 のテストデータに対する精度を示す．平均正答率は 0.99 以上の非常に高い値を示した．これらのモデルについても平均再現率に着目し比較すると，モデル B1 が 0.94，モデル B2 が 0.927，モデル B3 が 0.955 となり，この場合も機械学習手法に LightGBM を用いたモデル B3 が最も良い結果を示した．

モデル A とモデル B の結果を比較すると，ランダムに漏水点を配置したモデル A よりも，既往の管種ごとの漏水実績を反映させたモデル B の再現率が高い．このことから，精度の高い漏水予測モデルの構築には，既往の漏水実績を分析し，管種ごとの漏水のしやすさを踏まえた予測とすることが非常に有用であると考えられる．本研究では，千葉県営水道に対するヒアリング調査に基づき，管種が 2 種類しか存在しない簡略化した管網を想定としている．上水道管路の管種は様々である [9]ので，今後はより詳細な漏水実績に関する分析が必要である．

(6) スマートメータの導入率を変化させた場合

本研究では，地域内のスマートメータ導入率が低い場合を想定し，メータの導入率を 50%，25%，10%と変化させ，漏水予測モデルの精度評価を行った．具体的には，これまでの検討で最も精度の高かった LightGBM を用いたモデル B3 について，メータの導入率を 50%，25%，10%と変化させ，ランダムにメータ設置点を設定した．

表-6 に結果を示す．平均正答率はどのモデルにおいても 0.99 以上の非常に高い値を示した．平均再現率に着目し比較すると，説明変数に用いるスマートメータの割合が小さくなるにつれて再現率が低くなったものの，メータの圧力情報が全体の 10%の場合でも平均再現率 0.92 と高い精度を示した．

表-6 スマートメータの導入率を変化させた場合の予測精度

	モデル B3		
導入率	50%	25%	10%
平均正答率	0.999	0.999	0.999
平均適合率	0.964	0.961	0.939
平均再現率	0.946	0.938	0.92
平均 F 値	0.955	0.949	0.93

5. まとめ

本研究では，漏水位置の早期発見を目的として，スマートメータから得られる情報を活用し，機械学習による漏水予測モデルの構築を行った．

漏水検知モデルでは，メータ導入率を変えたどのモデルでも再現率は 0.9 と高い値を示した．また，スマートメータの導入率が低い場合でも漏水を検知できる可能性はあるが，10%の導入率では誤検知も多い結果となった．

漏水位置予測モデルでは，漏水点をランダムに設定したデータセット A，漏水点を管種ごとの漏水率に応じて設定したデータセット B のいずれのモデルにおいても，機械学習手法に LightGBM を用いたモデルが最も精度が良いものとなった．また，漏水シナリオが現実に近いデータセット B では，スマートメータの導入率が低い場合でも，漏水点を予測できる可能性が示された．

参考文献

1) 東京都水道局：東京の漏水防止，令和 3 年度版
2) 高桑哲男：配水ネットワークの解析法とその応用，電気学会雑誌，Vol. 966，No. 6，pp. 527-534，1976.
3) 土木学会：水理公式集，p. 511，2018.
4) 生活知恵袋ホームページ，洗濯機の水使用量，http://www.saikatu-cb.com/kiwami/siyous01.html (参照 2023-6-25)
5) Markus M. B. and Hans P. K. and Raymond T. and Ng, J. S. : LOF: identifying dencity-baced local outliers, Proc. ACM SIGMOID Record, Vol.29, No.2, pp. 93-104, 2000.
6) Breiman, L.: Random Forests, Machine Learning, Vol. 24, No. 2, pp. 123-140, 1996.
7) Chen, T. and Guestrin, C.: XGBoost: A scalable tree boosting system, Advances in Pure Mathematics, Vol. 6, No. 9, pp. 785-794, 2016.
8) Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wie Chen, Weidong Ma, Qiwei Ye, Tie-Yan Liu: LightGBM: A Highly Efficient Gradient Boosting Decision Tree, In Advances in Neural Information Processing Systems, pp. 3149-3157, 2017.
9) 丸山喜久，永田茂，若松加寿江：東北地方太平洋沖地震における上水道管路被害分析に基づく広域被害予測に関する検討，日本地震工学会論文集，Vol.15，No.7，pp.7_416-7_417，2015.