

D-020

斜面防災システムにおけるセンサのカテゴリライズを用いた運用コストの削減

Reducing Maintenance Cost of Mud Slide Detection Sensor Network with Categorization of sensor data

加藤 和弥[†]
Kazuya kato

原田 史子[‡]
Fumiko Harada

島川 博光[‡]
Hiromitsu Shimakawa

1. はじめに

我が国では台風や集中豪雨などによる土砂災害が深刻な被害を与えている。中でも斜面崩壊による被害者数は最大比率を持つ [1]。

その最も重要な対策は、斜面崩壊の予兆を事前に把握し、適切な警告を発令することで人的・物的被害を未然に防ぐことである。斜面防災システム [2] では、無線センサネットワークを構築して地盤内のデータを収集し、斜面崩壊の危険性を予測している。しかしながらこのシステムには保守管理業務などによる人的負担の面で問題がある。

本研究では、同じような特性を持つデータを同定し各点に相当するセンサデータをカテゴリライズすることで、斜面防災システムの人的負担を削減する手法を提案する。提案手法ではカテゴリ構成ごとに信頼度の指標を、相関係数に基づいてユーザに提示する。ユーザはこれを元にカテゴリ構成を決定する。カテゴリ構成が定めれば、「処理すべきデータ数の減少」と「システムの故障時の対応頻度削減」により、人的負担を削減できる。

提案手法で有効なカテゴリ構成が導出できるか検証した。結果として、処理すべきデータ数が 60.56% になり、システム故障時の対応頻度が 40.24% となった。かつ、カテゴリライズの結果の信頼性を示す指標である相関係数は、0.67 と高い値を得られた。

2. 斜面防災システムの課題

2.1 斜面防災システム

我々は、斜面崩壊の兆候を事前検知し近隣住民を安全に避難させることを目的とした斜面防災システムを構築している [2]。本システムは無線センサノードを斜面上に設置し、無線センサネットワークにより取得されたセンサ情報をリアルタイムで分析することで、斜面崩壊の兆候を検知している。その構成を図 1 に示す。

本システムでは各無線センサノードに接続された複数のテンシオメータから取得した斜面上の多点・多深度での間隙水圧と、基地局に有線接続された気象センサが取

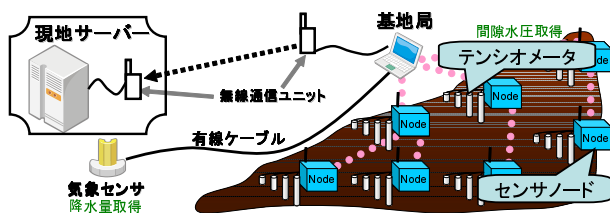


図 1: 斜面防災システム [2] の構成図

得した降水量を、崩壊兆候の分析に利用する。間隙水圧は土砂の隙間の地下水による圧力である。降雨が発生し、間隙水圧が上昇すると、斜面崩壊の危険性が増す。

2.2 斜面防災システムの運用コスト

現状の斜面防災システムでは、以下の3つの要因による人的負担の大きさが問題になっている。

データ分析コスト：現在斜面崩壊の兆候検知アルゴリズムのパラメータ調整のための、各テンシオメータから取得されたデータの手動分析による人的負担。

定期的な保守管理コスト：本システムのセンサノードは電池駆動であるため、定期的に電池交換作業などの保守管理が必要であり、人的負担がかかる。

緊急的な保守管理コスト：センサノードの不具合発生時、管理者が現場に赴き早急に対処する人的負担がある。この修復作業は管理者にとって想定外の作業であり、昼夜を問わず現場に赴く必要があるため人的負担は非常に大きいと言える。

以上のコストを、本研究では運用コストと定義する。

斜面防災システムは、将来的には交通量の多い道路に面した大規模斜面に適用され、センサネットワークも大規模になる。現状では、センサノードの数だけ運用コストが増大するため、大規模斜面に対応することは難しい。

2.3 既存研究

斜面防災システムのセンサネットワークは、ある間隔で配置されるセンサノードにより構築されている。センサノードの配置、センサノード間の通信間隔、通信経路によりセンサネットワークの耐故障性は変動する。耐故障性を向上させる既存手法として、センサノードの電池残量に応じて最適な通信経路を導出することでセンサネットワークの稼働時間を延長させる手法 [4] や、事前にセンサノードの最適な配置をシミュレートすることで、センサネットワーク全体の耐故障性を向上させる手法 [5] が研究されている。耐故障性向上により緊急的あるいは定期的な保守管理の頻度が減らせ、運用コストを削減できる。

しかし、斜面防災システムが使用している間隙水圧値は地盤の形状・性質に強く影響されるため、事前にその特性を把握できない。そのため、既存研究 [4, 5] では冗長なセンサノードがないように配置することができない。ノード配置の冗長性に因り、取得されるセンサデータは冗長性を含むことになる。冗長性のあるデータを処理することは運用コストの増大に繋がる。そのため、既存研究 [4, 5] では、十分に運用コストを削減できておらず、大規模斜面に適応するのは難しい。よって、これらの研究による耐故障性向上に加えて、データの冗長性も考慮した運用コスト削減が必要である。

[†]立命館大学大学院 理工学研究科

[‡]立命館大学 情報理工学部

3. カテゴリライズによる運用コスト削減

3.1 カテゴリライズ手法の概要

本研究では斜面崩壊検知システムの検知性能を維持しながら冗長性のあるデータ処理による運用コストの削減も行える手法を提案する。

各テンシオメータから得られるセンサデータには似た特性を持つデータ群も存在する。このような冗長なデータを同定しテンシオメータが設定されている観測点をカテゴリライズすることで、テンシオメータの代替やカテゴリごとの代表データ選出ができる。これにより「データ分析コスト」と「緊急的な保守管理コスト」の2つの運用コストが削減できる。

本手法は、土壌雨量と間隙水圧の関係を各観測点の地質の特性として同定し、それらを指標として観測点をカテゴリライズすることで冗長なデータを検出する。

土壌雨量とは土壌中の水分量である。ある地点の土壌雨量はタンクモデル [3] に基づき算出する。土壌雨量と間隙水圧の関係とは、ある地点における降雨毎の土壌雨量に対する間隙水圧の変化である。特性は土壌の性質による雨水の浸透しやすさ由来する。本研究では、この特性を水分特性曲線 [7] と呼ぶ。

カテゴリライズの概要を図2に示す。各テンシオメータに対応する水分特性曲線を過去のセンサデータから導出する。テンシオメータを使って異なる観測点から収集されたデータ間の水分特性曲線の類似性を相関係数として求め、与えられた閾値以上の相関係数を持つ観測点同士を同じカテゴリとするように、観測点群をいくつかのカテゴリに分割する。ユーザ入力として得られる閾値により、カテゴリ構成と斜面崩壊検知の信頼性は変化する。ユーザは閾値に対するカテゴリ構成と崩壊の兆候検知の信頼性の変化を参照し、カテゴリ構成を決定する。以下にカテゴリライズのプロセスの詳細を述べる。

3.2 水分特性曲線の導出

水分特性曲線の導出法を以下に示す。

1. テンシオメータから得られる間隙水圧と土壌雨量の値を二次元平面上にプロットし散布図を得る。得られた散布図は図3に示すような形状を持つ。
2. 1. で得られた散布図から、図3に示すような近似曲線を求め、これを水分特性曲線とする。水分特性曲線はヒステリシスを持つ。

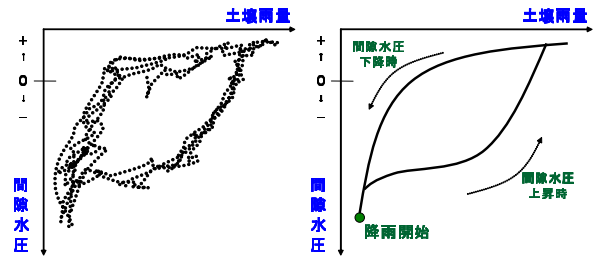


図3: センサデータを用いた散布図と水分特性曲線

これにより、水分特性曲線が導出できる。水分特性曲線は各テンシオメータの特性を表しているため、これを指標としてカテゴリライズするが、この状態では、曲線間の類似性を数値化することが難しい。そこで、各テンシオメータが収集した土壌雨量、間隙水圧それぞれを土壌雨量ベクトル $D = (d_1 \ d_2 \ d_3 \ \dots \ d_n)$ 、間隙水圧ベクトル $P = (p_1 \ p_2 \ p_3 \ \dots \ p_n)$ とベクトルで表す。ベクトル D とベクトル P は土壌雨量、間隙水圧の各値を時系列順に並べることで得られる。2つのテンシオメータの土壌雨量ベクトルと間隙水圧ベクトルから、水分特性曲線間の類似性を相関係数として導出する。

3.3 カテゴリライズ指標による階層的クラスタリング

各テンシオメータ T_i の水分特性曲線を、間隙水圧ベクトル P_i と土壌雨量ベクトル D_i の組 $(P_i \ D_i)$ で表せる。テンシオメータ群 $\{T_1, \dots, T_n\}$ の両ベクトルの組の集合 $N = \{(P_i, D_i), \dots, (P_n, D_n)\}$ を階層的クラスタリング [9] によりクラスタリングする。

クラスタリングにおいて、テンシオメータ T_i, T_j 間の類似度を式 (2) で得られる相関係数で表す。また、クラスタリングの各ステップにおいて、 T_i, T_j に対応するデータが新たに1つのクラスタ T_{ij} として結合される時、 T_{ij} に対応するデータを式 (3) として新たに得る。このように以降のステップでは、クラスタ T_{ij} を式 (3) で与えられる間隙水圧ベクトル P_{ij} と土壌雨量ベクトル D_{ij} をもつ1テンシオメータとみなして、テンシオメータ群のクラスタリングを行う。

$$c = \frac{\sum_{j=1}^{n-1} (\sqrt{P_i^2 + D_i^2} - \overline{PD}_i) (\sqrt{P_j^2 + D_j^2} - \overline{PD}_j)}{\sqrt{\sum_{i=1}^n (\sqrt{P_i^2 + D_i^2} - \overline{PD}_i)^2} \sqrt{\sum_{j=1}^n (\sqrt{P_j^2 + D_j^2} - \overline{PD}_j)^2}} \quad (1)$$

$$\overline{PD}_i = \frac{\sum_{i=1}^n \sqrt{P_i^2 + D_i^2}}{n} \quad \overline{PD}_j = \frac{\sum_{j=1}^n \sqrt{P_j^2 + D_j^2}}{n} \quad (2)$$

$$(P_{ab}, D_{ab}) = \left(\frac{(P_a + P_b)}{2}, \frac{(D_a + D_b)}{2} \right) \quad (3)$$

3.4 ユーザ操作によるカテゴリ構成

同一カテゴリに属するテンシオメータが収集するデータ間の相関係数が小さい場合、斜面崩壊検知の信頼性が低くなる。逆に、同一カテゴリに属するテンシオメータが収集するデータ間の相関係数が大きい場合は、検知に対しての信頼性は高くなるが、相対的にカテゴリ数が多くなり、運用コストの削減量が低下してしまう。

最適なカテゴリ構成を確定させるためには、ユーザ入力とユーザによる判断が必要である。階層的クラスタリングの結果は、図4のようにデンドログラムで可視化できる。このデンドログラムは横軸に各テンシオメータが

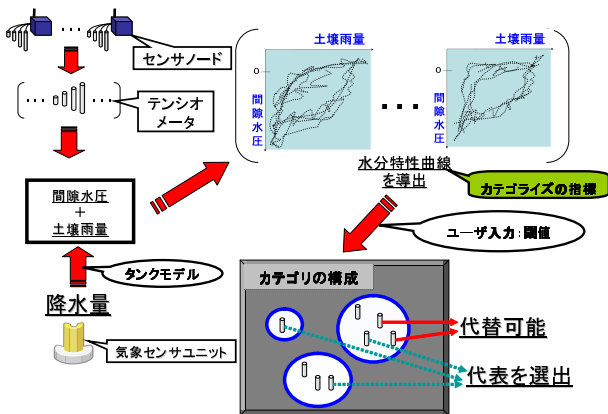


図2: カテゴリライズの流れ

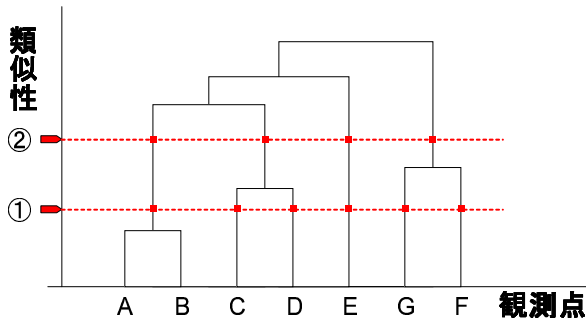


図 4: デンドログラム

設置された観測点を配置し、縦軸は類似度となる相関係数を示す。

ユーザは類似度を閾値として指定することで、その閾値におけるカテゴリ構成が得られる。図 4 を用いて例を示す。閾値を①の点に指定すると、観測点 A と B がひとつのカテゴリとなるため、カテゴリ数は 6 となる。②の点を指定すると、C と D、F と G が同じカテゴリとなるため、カテゴリ数は 4 になる。ユーザが閾値を指定し、そのカテゴリ構成を吟味することで、ユーザが許容できるカテゴリ構成が得られる。ユーザがカテゴリ構成を決定する流れを述べる。

1. ユーザが閾値を入力する。
2. デンドログラムから閾値に等しい相関係数に対応するカテゴリ構成と最低相関係数を示す。最低相関係数については 3.5 で後述する。
3. ユーザは結果を確認し、許容できるなら終了。そうでなければ 1. に戻る。

この 1. から 3. を繰り返すことにより、ユーザは自らが許容できるカテゴリ構成を得られる。

カテゴリ化により得たカテゴリ構成を用いて運用コストを削減することで、斜面崩壊検知の信頼性は低下する。最低相関係数は、信頼性が低下したとしても、最低限この値の示す信頼性を確保していることを示す。

3.5 最低相関係数の導出

カテゴリ化によりひとつのカテゴリとして定義した観測点のデータ群は冗長性を持った同一のデータとして処理する。実際には異なった観測点のデータを同一のデータとして処理するため、斜面崩壊検知の信頼性が低下する。

信頼性の低下の指標を相関係数に基づいて与える。相関係数として数値として示すことで、ユーザがカテゴリ構成ごとの信頼性を比較できるようになる。この指標となる相関係数を最低相関係数と定義する。

あるカテゴリ構成における最低相関係数を C_{min} とする。 C_{min} の導出手順を示す。このカテゴリ構成は、 M 個のカテゴリを有しているとする。それぞれのカテゴリごとに、カテゴリ内の観測点のデータ間の相関係数をすべての観測点 2 点の組み合わせで式 (1) により計算し、その最小値を求める。カテゴリごとの最小相関係数の集合 C が得られる。集合 C での最小相関係数を、このカテゴリ構成における最低相関係数 C_{min} とする。

3.6 カテゴリ構成による運用コスト削減

ユーザ操作により得られたカテゴリ構成を用いて、以下の 2 つの運用コストが削減できる。

データ分析コスト： テンシオメータをカテゴリ化することで、各カテゴリの代表となる観測点のデータを選出できる。あるカテゴリのデータ分析を代表となる観測点のデータのみで行うことで、データ分析コストを削減できる。

緊急的な保守管理コスト： 観測点をカテゴリ化することで、センサノードやテンシオメータの故障によりあるテンシオメータからデータが取得できなくても、同一カテゴリの他の観測点に設置されたテンシオメータのデータで代替できる。そのため、センサノードやテンシオメータが故障しても、システムの安全性に問題が生じるほど観測点が少なくならないうえに、修復作業を行わなくても済む。これにより修復作業に向かう頻度を減らすことができるため、緊急的な保守管理コストを削減できる。

4. 運用コスト削減手法の検証

4.1 実験目的

提案手法の有効性を検証する実験を実データを用いて行った。本実験の目的は、「提案手法を用いて運用コストの削減が可能なカテゴリ構成が導出できるかの検証」と、「カテゴリ構成と閾値の変化による運用コストの削減率変化の傾向分析」の 2 点である。

本実験においては、任意の閾値に対するカテゴリ化の効果を知るためのシミュレータを作成し、ある閾値に対するカテゴリがもたらすコスト削減量と信頼性の計算を自動化した。

4.2 実験内容

斜面防災システムによって、2009 年 6 月から 11 月にかけて、降水量が継続的に観測できた 6 観測日分、7 月 10 日、7 月 13 日、8 月 8 日、9 月 12 日、11 月 1 日、11 月 17 日の前後の観測日の降水量と間隙水圧のデータを用いる。

各観測日において最大 30 個のテンシオメータのデータがある。各観測日に対し、あるテンシオメータのデータが数分から数時間同値であるか、データ損失が 5 割を超える場合、そのテンシオメータは故障しているとみなしその観測日データから除外する。

各観測日 Day_i に対して、その前の観測日 Day_{i-1} のデータを用いて提案手法を適用し、各閾値に対応するカテゴリ構成を得る。ただし観測日 7 月 10 日に対しては、7 月 10 日のデータに提案手法を適用する。当該カテゴリ構成を観測日 Day_i のセンサデータに適用した場合のデータ分析コスト削減率と緊急的な保守管理コストの削減率を導出する。

観測日 Day_{i-1} のデータとユーザ入力としての閾値 ϵ から、カテゴリ数 K と、各カテゴリに属するテンシオメータ集合の $M = \{M_1, \dots, M_K\}$ が得られたとする。また、観測日 Day_i のデータが取得されたテンシオメータ数を M_t とする。分析コストと緊急的な保守管理コストの削減率 C_d, C_c をそれぞれ式 (4)(5) に従い計算する。

$$C_d = \frac{(M_T - K)}{M_T} \times 100 \quad [\%] \quad (4)$$

$$C_c = \frac{(\text{集合 } M \text{ に含まれるデータ数が 2 個以上のクラスタ数})}{K} \times 100 \quad [\%] \quad (5)$$

4.3 検証結果

カテゴリ構成から運用コストの削減量を導出する。閾値を0.1から0.9まで変化させることで、運用コスト削減量の推移を確認する。データ分析コストの削減量を図5に、緊急的な保守管理コストの削減量を図6に提示する。

図5と図6より、閾値を大きくすることで、運用コストの削減量は小さくなる傾向が確認できる。詳しく見ていくと、データ分析コストについては、閾値0.3以下では、期間ごとの運用コスト削減量の差がほとんど見られない。そして、緊急的な保守管理コストにおいては、閾値0.7以下にて、運用コスト削減量の差がほとんどなく、閾値0.7以上に指定すると、急激にコスト削減量が低下する傾向が確認できる。よって、緊急的な保守管理コストについては、0.7以下の閾値を指定することで削減量を大幅に増やすことができると言える。

このように、ユーザ入力による閾値の指定により運用コストの削減量を得ることができる。

今回は、ユーザが閾値0.8を用いた場合の運用コスト削減量と最低相関係数の関係が最適であると指定したとする。表1と図7にて、閾値0.8における、運用コストの削減量を提示する。

各期間の平均を見ると、データ分析コスト C_d が60.56%に、緊急的な保守管理コスト C_c が40.24%におさえられる。そしてこのカテゴリ構成の信頼度としての最低相関係数は0.67であるということが確認できる。閾値を0.8として高く設定したが、それでも非常に高い値を得ることができた。以上より、閾値0.8において妥当な量の運用コストが削減可能なカテゴリ構成が導出できることを示した。

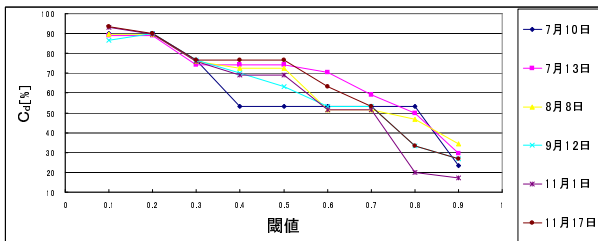


図 5: データ分析コストの推移

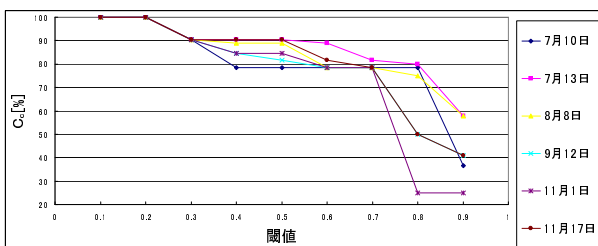


図 6: 緊急的な保守管理コストの推移

表 1: 閾値 0.8 における運用コスト削減量と最低相関係数

閾値: 0.8	7/10	7/13	8/8	9/12	11/1	11/17	平均
C_d	46.67	50.00	53.33	66.67	80.00	66.67	60.56
C_c	21.43	20.00	25.00	50.00	75.00	50.00	40.24
最低相関係数	0.775	0.887	0.446	0.670	0.445	0.768	0.67

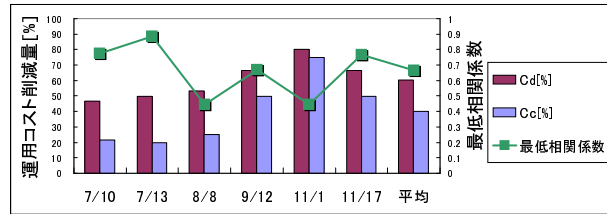


図 7: 閾値 0.8 における運用コスト削減量と最低相関係数

5. おわりに

本論文では、斜面防災システムにおける運用コストを削減するために、データの類似性に着目した観測点のカテゴリ分け手法を提案した。本手法は、土壌雨量と間隙水圧の関係に基づいてセンサが設置されている観測点をカテゴリ分けし、冗長なデータを検出することで運用コストを削減する。そして、ユーザ操作により最適なカテゴリ構成を確定させる。

実データを用いて、提案手法から運用コストが削減可能なカテゴリ構成が導出できるかを検証した。ユーザが入力とした閾値を相関係数0.8としたとき、データ分析コスト C_d が60.56%に、緊急的な保守管理コスト C_c が40.24%に削減できた。そして、最低相関係数は0.67と非常に高い値を得ることができた。これにより、提案手法の有用性が認められた。

今後は、階層的クラスタリング結果を過去のデータの変遷から導出することで、より現状の斜面防災システムのデータ群に則したカテゴリ構成を得られるようにし、本手法の信頼性を向上する。

参考文献

- [1] 牛山素行, 2004~2007年の豪雨災害による人的被害の原因分析, 河川技術論文集, Vol.14, pp.175-180. Jun. 2008.
- [2] 深川良一, センサーネットワークを利用した次世代型斜面防災システムの構築, 立命館大学 総合理工学研究機構 防災システム研究センター, 2008.
- [3] 岡田憲治, 牧原康隆, 新保明彦, 永田和彦, 国次雅司, 齊藤清, 土壌雨量指数, Tenki, vol.48, no.5, pp.349-356, 2000.
- [4] 内藤克浩, 他, 電源残量に基づくセンサネットワーク用ルーティングプロトコルの研究, 電子情報通信学会技術研究報告, vol.107, no.37, pp.49-54, 2007.
- [5] 斎藤洋, 南裕也, センサネットワークの性能尺度とその評価, 電子情報通信学会技術研究報告, vol.102, no.693, pp.129-132, 2003.
- [6] Y. Hayashida, Y. Kohgo, and I. Asano, A hysteresis model with the continuity of tangential slopes of soil water retention curves, Proceedings of the 2nd Asian Conference on Unsaturated Soils, UNSAT-ASIA 2003, Osaka, pp.413-418, 2003.
- [7] 乾湿繰り返しを考慮した水分特性曲線モデル, 東京農工大学 環境地盤研究室, http://www.tuat.ac.jp/~kohgo/soilw.html.
- [8] 西田義親, 八木則男, 二木幹夫, 雨水浸透による地盤内発生間隙水圧と斜面安定, 土質工学会論文報告集, 19(1) xiii, 1979/03.
- [9] 斎藤堯幸, 宿久洋 著, 関連性データの解析法, 共立出版, 2006.