

半教師あり学習を用いた GMM による 異常放射からの地震前兆識別

曾我 昌士 酒向 慎司 北村 正

名古屋工業大学

1 はじめに

日本は地震大国として知られている。地震被害を軽減するためには発生を事前に検知することも有効である。地震発生の数日から数週間前に環境電磁波の異常放射が観測されることが知られているため、これを検出して地震の発生を事前に捉える研究が行われている。

先行研究では、まず環境電磁波の定常的な振る舞いから外れたものを隠れマルコフモデル (HMM) を用いて異常として検出する。次に異常の要因は多数存在しているため、検出された異常に対して混合ガウスモデル (GMM) を用いて地震前兆とその他の異常を識別する枠組みが提案されている [1]。この枠組みでの GMM は地震前兆 (正例) とその他の異常 (負例) のラベルありデータで学習されるが、正例クラスの学習データは大地震の事例の少なさから十分でないと考えられる。そのためモデルを詳細に学習できず、識別誤りが多く見られた。そこで本研究では、ラベルのないデータも GMM の学習に使用し、モデルを詳細化することで識別精度の向上を図る。

2 環境電磁波からの異常放射検出

環境電磁波には 1 日の周期性がある。これをモデル化するため、1 日あたり 24 フレームの特徴量を抽出して 1 日分の特徴ベクトル系列を作成し、HMM を学習する。

異常検出も、テストデータについて特徴ベクトル系列を作成して学習されたモデルに対する尤度を計算し、式 (1) を満たすものを異常とする。 μ と σ はテストデータの尤度の平均と分散であり、 k は定数である。

$$Likelihood \leq \mu - k * \sigma \quad (1)$$

3 異常放射からの地震前兆識別

正例クラスの学習データは十分でないと考えられるため、過去の観測データをラベルなしデータとして学習データに追加し、半教師あり学習を行った GMM [2] で地震前兆の識別を行う。

学習データとは別に地震の発生が確認されるデータ群を用意し、GMM の学習データの選別を自動化するために、HMM と GMM を利用してラベルなしデータの選出とラベル付けを行う。まず、このデータ群について HMM に対する尤度を求める。このデータ群から、テストデータに対して異常判定を行った閾値を下回るものをラベルなしデータとして選出することで、テストデータと同程度の異常データを集めることができる。次に、ラベルありデータのみで学習した GMM を初期モデルとして、現在の GMM を用いてラベルなしデータが各クラスに識別される確率を計算してラベルなしデータにラベルとして付加するステップと、その確率とラベルなし

表 1 実験結果

k		Precision	Recall	F 値
1.00	従来法	0.138	1.000	0.242
	提案法	0.150	0.750	0.250
1.25	従来法	0.105	0.500	0.174
	提案法	0.111	0.750	0.194
1.50	従来法	0.167	0.500	0.250
	提案法	0.176	0.750	0.286

データ、ラベルありデータを用いて GMM パラメータを再推定するステップをくり返すことでパラメータを更新する。半教師あり学習 GMM の混合数はベイズ情報量規準 (BIC) を最小化するものを選択する。

4 評価実験

宮城県栗原市で観測された 5 年分のデータを学習データ、1 年分のデータをラベルなしデータ、2008 年のデータをテストデータとし、 k を変化させて評価実験を行った。検出対象は 2008 年に発生した 4 件の大地震 (地震エネルギー 12 以上) とした。評価には F 値を用いる。地震は前日からその 2 週間前までに 1 日でも前兆と識別されれば検出できるものとする。

5 結果と考察

実験結果を表 1 に示す。わずかではあるが提案法による F 値の向上が確認できる。 k を大きくすることで F 値の向上も大きくなっているが、 $k = 1.25, 1.50$ では異常度の判定の時点で 1 件の地震の前兆を見落とす結果となったため、ただ k を大きくするだけではなく地震前兆を見落とさないような尤度を設定する必要がある。

6 むすび

本研究では、ラベルなしデータを用いた半教師あり学習による GMM の詳細化を行った。評価実験からは、 F 値の向上が確認できた。今後の課題として、地震を見落とさない範囲で低い閾値を設定できる方法の検討などが挙げられる。

参考文献

- [1] 中村 祐二郎, 酒向 慎司, 北村 正, “HMM による環境電磁波の異常検知と地震前兆信号の判別,” 2013 信学全大, 情報・システム講演論文集 2, no.D-12-9, pp.102, Mar. 2013.
- [2] Xiaojin Zhu, “Semi-Supervised Learning Tutorial,” 2007 ICML Tutorial Schedule, <http://oregonstate.edu/conferences/event/icml2007/tutorials.html>, 参照 Aug.1,2013.