

의사 샘플 신경망을 이용한 토석류 퇴적 모델의 파라미터 추정

허경용*, 이창우**, 박충식***

Parameter Estimation in Debris Flow Deposition Model Using Pseudo Sample Neural Network

Gyeongyong Heo *, Chang-Woo Lee **, Choong-Shik Park ***

요약

토석류 퇴적 모델은 토석류에 의한 피해지 예측을 위해 random walk model(RWM)을 사용하여 구성된 모델로 피해지 예측에서 그 효용성이 입증되었지만 몇 개의 자유 파라미터가 실험적으로 결정되어야 하는 문제점이 있다. 파라미터를 자동으로 추정하기 위한 방법은 여러 가지가 있지만 토석류 데이터는 학습 데이터의 크기가 작아 기존 학습 기법을 적용하는데 어려움이 있다. 이 논문에서는 학습 데이터 크기 문제를 완화할 수 있는 신경망의 변형인 의사 샘플 신경망을 제안하였다. 의사 샘플 신경망은 기존 샘플로부터 의사 샘플을 생성하고 이를 학습에 사용한다. 의사 샘플은 해공간을 평탄화시키고 극부 최적해에 빠질 확률을 줄여줌으로써 기존 신경망에 비해 안정적인 파라미터 추정이 가능해진다. 이러한 사실은 실험 결과 통해 확인할 수 있다.

▶ Keywords : 파라미터 추정, 의사 샘플, 신경망, 토석류 퇴적 모델

Abstract

Debris flow deposition model is a model to predict affected areas by debris flow and random walk model (RWM) was used to build the model. Although the model was proved to be effective in the prediction of affected areas, the model has several free parameters decided experimentally. There are several well-known methods to estimate parameters, however, they cannot be applied directly to the debris flow problem due to the small size of training data. In this paper, a modified neural network, called pseudo sample neural network (PSNN), was

•제1저자 : 허경용 •교신저자 : 박충식

•투고일 : 2012. 08. 17, 심사일 : 2012. 10. 02, 게재확정일 : 2012. 10. 18.

•이 논문은 2012년 한국컴퓨터정보학회 제46차 하계학술대회에서 발표한 논문("인공 신경망을 이용한 토석류 퇴적 모델 파라미터 추정")을 확장한 것임.

* 동의대학교 전자공학과(Dept. of Electronic Engineering, Dong-Eui University)

** 국립산림과학원 산림방재연구과(Division of Forest Management, Korea Forest Research Institute)

*** 영동대학교 스마트IT학부(Dept. of Smart Information Technology, Youngdong University)

proposed to overcome the sample size problem. In the training phase, PSNN uses pseudo samples, which are generated using the existing samples. The pseudo samples smooth the solution space and reduce the probability of falling into a local optimum. As a result, PSNN can estimate parameter more robustly than traditional neural networks do. All of these can be proved through the experiments using artificial and real data sets.

▶ Keywords : Parameter Estimation, Pseudo Sample, Neural Network, Debris Flow Deposition Model

I. 서 론

우리나라의 토석류 피해는 대부분 집중 호우에 의해 발생되고 그 위치가 국지적으로 밀집된 양상을 보인다. 따라서 토석류 피해 위험지를 예측하고 이를 통해 사방구조물의 적지 선정 및 경계 피난 체제 구축을 통해 피해의 규모를 줄일 수 있을 것으로 기대할 수 있다. 위험지 예측 및 구조물 위치 선정을 위해 이전 연구에서는 random walk model(RWM)[1]을 이용하여 토석류 퇴적 모델을 구축하고 이를 통해 성공적인 결과를 얻을 수 있었다[2]. 하지만 RWM 모델에는 최적화시켜야 할 세 개의 파라미터 값이 존재하며 이 값들은 경험적으로 결정되었다. 보다 넓은 지역에 토석류 모델을 적용하기 위해서는 파라미터 값을 자동으로 결정할 수 있는 방법이 필요하다.

파라미터 값을 추정하는 방법은 여러 가지가 있지만, 토석류 데이터의 경우 기존의 방법을 적용하기 위한 충분한 데이터를 확보하는데 어려움이 있다. 이처럼 학습 샘플이 적은 경우 학습된 시스템의 신뢰성이 떨어지는 문제를 SSS (Small Sample Size) 문제라고 한다[3]. 이 논문에서는 인공 신경망을 사용함에 있어 SSS 문제를 완화하기 위해 의사 샘플을 생성하고 이를 학습에 사용하는 의사 샘플 신경망 (PSNN, Pseudo Sample Neural Network)을 제안한다. 제안한 PSNN은 의사 샘플을 통해 국부 최적해에 수렴하는 현상을 줄일 수 있으며 따라서 보다 안정적인 파라미터 추정이 가능해진다. 이러한 사실들은 실험 결과를 통해 확인할 수 있다.

이 논문의 구성은 다음과 같다. 2장에서는 토석류 퇴적 모델과 토석류 퇴적 모델의 파라미터 추정에서의 요구 사항을 논한다. 3장에서는 제시된 요구 사항을 만족시킬 수 있는 신경망의 변형인 샘플 신경망을 제안하며 4장에서는 실험 결과를 통해 의사 샘플 신경망의 효용성을 보인다. 향후 연구 방향은 5장에서 제시한다.

II. 토석류 퇴적 모델 및 파라미터 추정

1. 토석류 퇴적 모델

우리나라 국토의 64%를 차지하는 산지는 매년 장마, 집중 호우, 태풍 등으로 인해 발생하는 산지 토사 재해(산사태, 토석류 등)로 몸살을 앓고 있다. 특히 토석류는 산사태로 생긴 토사와 계류바닥에 쌓인 흙, 돌, 바위 등이 물과 함께 흘러내려 주택, 농지를 덮침으로써 일어나는 재해로 우리나라 산지 계곡에서 발생하는 재해의 대부분을 차지하고 있다. 토석류로 인한 피해는 토석류의 흐름을 예측하고 이에 따라 적절한 위치에 구조물을 설치함으로써 줄일 수 있으며, random walk model (RWM)은 토석류의 퇴적 행태를 나타내기 적합한 모델이다[1]. RWM은 토석류가 하루로 내려올 때 경사가 급한 곳으로 진행된다는 가정 하에 이동방향 및 퇴적 위치를 결정하며 이를 위해 1회에 흘러 내리는 토사량 (1회 유송량), 토석류 이동시 유체의 직진성 (관성가중치), 토석류가 퇴적하는 경사도 (정지조건) 등 3개의 매개변수를 통해 토석류의 이동 및 피해범위를 예측한다. RWM을 2006년 토석류가 집중적으로 발생한 인제군을 대상으로 분석한 결과를 인제군의 다른 토석류 발생지 4개소에 적용해 본 결과 퇴적면적에 있어 56%의 일치율을 보였으며 이를 이용하여 토석류에 의한 위험 지도를 시범 제작하였다[2].

토석류 모델을 적용함에 있어 관건은 위의 세 가지 파라미터의 값을 결정하는 것이다. 이들 파라미터는 직접적으로 관찰할 수 없는 값들로 지역에 따라 다른 값을 가지며 이전 연구에서는 경험적으로 그 값이 결정되었다. 따라서 보다 광범위한 지역에 적용하기 위해서는 토석류 모델에 필요한 세 가지 파라미터를 자동으로 결정하는 방법이 필요하며 이 논문은 관측 가능한 값들로부터 RWM의 파라미터 값들을 자동으로 예측할 수 있도록 해주는 파라미터 추정 방법을 제시함에 그 목적이 있다.

2. 토석류 퇴적 모델을 위한 파라미터 추정

토석류 연구에서 관측 가능한 값들은 표고별 유역면적곡선과 유역면적별 계류차수의 두 가지이며, 직접적인 관측이 불가능하지만 RWM에서 필요로 하는 파라미터 값들은 1회 토사량, 관성 가중치, 정지조건의 세 가지이다. 이 논문에서는 이들 사이의 관계를 자동으로 학습하여 관측 가능한 값들로부터 RWM의 파라미터들을 자동으로 추정하는데 목적이 있다. 이러한 관계 학습을 위해 사용할 수 있는 방법은 많지만 토석류 연구에서는 몇 가지 제한 사항이 있다. 먼저 입력(관측 가능한 변수)과 출력(RWM 파라미터) 사이의 관계가 명확하게 알려져 있지 않다. 비록 입력과 출력 사이의 대략적인 연관성은 알려져 있지만 이들 사이의 정확한 관계를 수식으로 표현하는 것은 불가능하다. 또 다른 문제점은 현재 사용할 수 있는 데이터가 적다는 점이다. 현재 학습에 사용할 수 있는 데이터는 5개 지역의 데이터뿐이므로 일반적인 파라미터 추정 방법을 적용하는데 어려움이 있다. 끝으로 향후 지속적으로 데이터가 추가될 예정이므로 데이터가 추가되는 경우 추가된 데이터를 학습된 모델에 쉽게 반영할 수 있어야 한다. 토석류 퇴적 모델의 파라미터 추정을 위한 요구 사항은 다음과 같이 요약할 수 있다.

- ① 입력과 출력 사이의 관계를 명시적으로 표현할 수 없는 경우에 사용 가능하여야 한다.
- ② 데이터가 적은 경우에도 사용이 가능하여야 한다.
- ③ 데이터가 추가되는 경우 기존에 학습된 모델에 반영할 수 있어야 한다.

다른 파라미터 학습 방법에 비해 신경망은 입력과 출력 사이의 관계 자체를 학습할 수 있으며 데이터가 추가되는 경우 사용할 수 있는 점진적 학습 방법(incremental learning)이 정립되어 있다[4-6]. 하지만 신경망 역시 다른 학습 방법과 마찬가지로 학습을 위해서는 일정량 이상의 학습 데이터를 필요로 한다. 학습에 필요한 데이터의 크기에 대한 연구는 오랜 역사를 가지고 있지만[7-9] 여전히 풀어야 할 숙제로 남아있다. 가장 단순한 선형 분류기의 경우에도 데이터 차원의 2배 이상의 학습 샘플을 요구하는 것으로 알려져 있으며 이는 주어진 문제와 학습 모델 및 학습 방법에 영향을 받는다. 일반적으로 학습 샘플은 많을수록 유리하다. 하지만 모든 문제에서 충분한 양의 샘플을 획득할 수는 없으며 토석류 데이터 역시 마찬가지이다. 따라서 이 논문에서는 데이터의 부족을 해결하기 위해 의사 샘플(pseudo sample)을 이용하여 신경망을 학습시키는 방법을 제안한다. 제안하는 방법은 기존 샘플을 이용하여

의사 샘플을 생성하고 이를 신경망 학습에 사용함으로써 학습된 모델의 안정성을 높일 수 있으며 따라서 추정된 파라미터의 신뢰성을 높일 수 있다.

III. 의사 샘플 신경망

의사 샘플 신경망(Pseudo Sample Neural Network, PSNN)은 신경망의 변형으로 학습 샘플의 구성에서 차이점이 있고 학습 및 평가 방법은 신경망과 동일하다. 따라서 PSNN은 학습 과정에서 의사 샘플 구성을 위한 전처리 단계를 도입함으로써 구성할 수 있다. PSNN의 학습 방법은 토석류 데이터의 크기를 고려하여 일괄 학습(batch learning)을 사용하였다. PSNN에서 의사 샘플은 기존 샘플을 평균으로 가지는 가우스 분포에서 생성하며 이 때 가우스 분포의 분산은 샘플의 분산을 샘플의 수로 나눈 값을 사용하였다. 가우스 분포 이외의 분포를 활용할 수도 있지만 샘플의 수가 적은 경우에는 분포를 추정하기가 어려우며, 모든 분포는 가우스 분포의 혼합으로 설명할 수 있고 샘플의 수가 많은 경우 모든 분포는 가우스 분포로 수렴한다는 중심 극한 정리(central limit theorem)[10]에 따라 이 연구에서는 가우스 분포를 사용하였다. N 개의 샘플 $X = \{x_1, x_2, \dots, x_N\}$ 이 주어진 경우 의사 샘플 X_{PS} 는 N 개의 가우스 분포에서 생성되며 식 (1)과 같이 나타낼 수 있다.

$$X_{PS} = \bigcup_{i=1}^N \left\{ x_j' \mid x_j' \sim N\left(x_i, \frac{\Sigma}{N}\right), j=1, \dots, N_{PS} \right\} \quad (1)$$

이 때 Σ 는 샘플 X 의 분산-공분산 행렬을 나타내고, N_{PS} 는 각 분포에서 생성하는 샘플의 개수로 이 논문에서는 실험적으로 결정하였다. 생성된 의사 샘플 X_{PS} 는 의사 샘플 신경망의 학습에 사용한다.

먼저 제안하는 PSNN의 실효성을 이차 방정식의 학습을 통해 보인다. 그림 1은 2차 방정식을 학습하기 위해 -10에서 10까지 균등하게 11개의 x 값을 취하고 계산된 $y = \frac{1}{120}(x^2 + 2x - 1)$ 값에 가우시안 노이즈를 추가한 것이다. 그림 2은 그림 1의 샘플로 식 (1)을 이용하여 생성한 의사 샘플을 나타낸다.

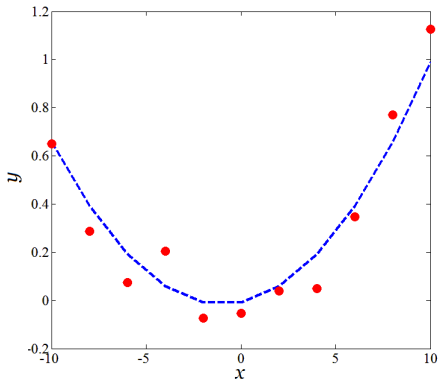


그림 1. 2차 방정식과 잡음이 섞인 샘플
Fig. 1. A quadratic equation and its noisy samples

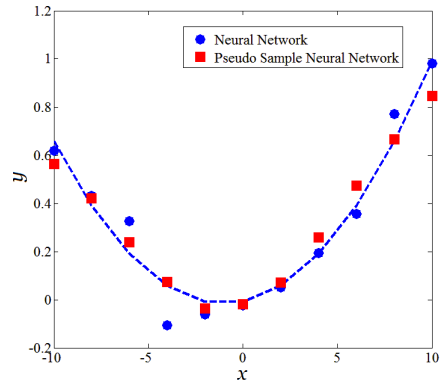


그림 3. 신경망과 의사 샘플 신경망의 학습 결과
Fig. 3. Training result of a traditional neural network and a pseudo sample neural network

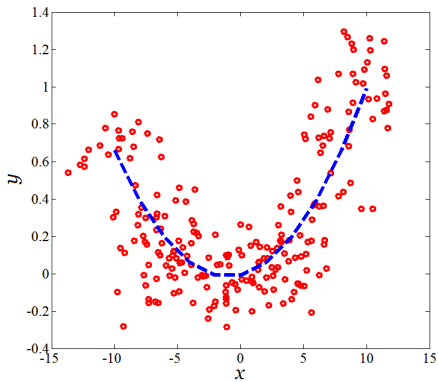


그림 2. 가우시안 분포를 이용하여 생성한 의사 샘플
Fig. 2. Pseudo samples using Gaussian distributions

X (그림 1)는 전통적인 신경망에서 학습 샘플로 사용되며 X_{PS} (그림 2)는 의사 샘플 신경망의 학습에 사용된다. 두 학습 샘플 집합 X 와 X_{PS} 를 이용하여 신경망을 학습시키고 학습된 신경망에서의 결과를 비교한 예가 그림 3이다. 그림 3에서 보아 알 수 있듯이 의사 샘플을 이용하여 학습한 신경망의 경우 보다 실제 함수에 근접한 결과를 보여주고 있다. 이 때 두 신경망의 비교는 실제 함수의 값(y_i)과 신경망에서 출력된 값(o_i)의 차이를 제공하여 합한 오류(sum square error)로 비교하였다.

$$E = \sum_{i=1}^N \|y_i - o_i\|^2 \quad (2)$$

표 1. 토석류 데이터
Table 1. Debris flow data

지역	출력 (RVM 파라미터)			입력 (관측 가능한 변수)	
	1회 토사량 (1~100m3)	관성가중치 (1~5)	점지조건 (0~15°)	표고별 유역면적곡선	유역면적별 계류차수
1(평창)	4.0	4	7.8	0.379	2.5785
2(인제)	5.1	2	4.2	0.330	2.3685
3(봉화)	6.3	3	6.7	0.366	2.5502
4(광양)	10.9	5	7.3	0.229	2.346
5(제천)	1	5	8.7	0.397	2.3362

2차 곡선의 학습을 100회 반복한 결과 신경망의 평균 오류는 0.071이며 의사 샘플 신경망의 평균 오류는 0.042였다. 의사 샘플을 통해 학습 샘플의 수가 늘어나면 신경망의 학습은 평탄화된 해 공간에서 진행되고 국부 최적해에 빠질 확률이 줄어든다. 따라서 위의 실험 결과에서 나타난 바와 같이 실제 곡선에 가까운 곡선을 학습할 확률이 증가하여 적은 오류를 보게 된다.

IV. 실험 결과

1. 인공 데이터를 이용한 실험

실험에서는 먼저 앞 장에서 사용하였던 2차 함수 학습의 예를 확장하여 신경망과 의사 샘플 신경망의 성능을 비교하였다. 앞 장에서는 2차 함수의 샘플을 얻기 위해 y 값에만 가우시안 잡음을 첨가한데 비해 이 장에서는 x 값에도 가우시안 잡음을 첨가하였다. 또한 은닉층의 개수와 각 은닉층의 노드 개수를 변화시킨 서로 다른 구조의 신경망에 대해 실험하여 구조에 관계없이 의사 샘플 신경망이 신경망에 비해 안정적인 결과를 보인다.

그림 4는 노드 개수에 따른 평균 오류를 나타낸다. 이 때 노드의 개수는 전체 은닉층에 있는 노드의 수를 합한 것으로 노드의 개수가 증가함에 따라 은닉층의 개수도 증가한다. 하나의 구조에 대해 50회 반복 실험하였고 오류를 평균하여 그림 4에 나타내었다.

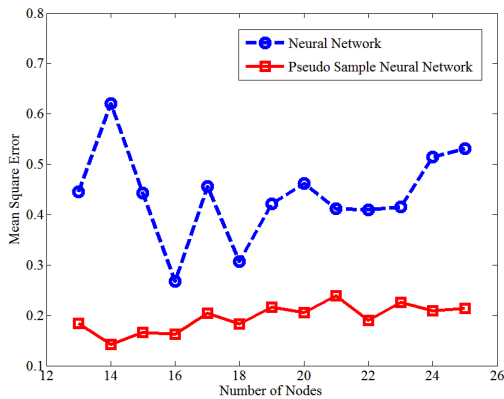


그림 4. 은닉층 노드 개수에 따른 자승 평균 오류
Fig. 4. Mean square errors with respect to the number of nodes in hidden layers

그림 4에서 보아 알 수 있듯이 의사 샘플을 이용한 신경망의 오류가 일반 신경망에 비해 작다. 이는 학습 샘플의 증가에 따라 해공간이 평탄화된 때문으로 잡음 감소 효과로 이해할 수 있다. 노드 개수가 증가함에 따라 오류가 증가하는 것은 학습하고자 하는 모델이 간단한 2차 방정식으로 복잡한 구조의 신경망은 주어진 데이터에 과적응(over-fitting)되기 때문이다.

2. 토석류 데이터를 이용한 실험

제한한 의사 샘플 신경망은 토석류 퇴적 모델을 위한 파라미터 추정에 사용하였다. 실험에 사용한 데이터는 표 1과 같다. 표 1에서 입력은 표고별 유역면적곡선 값과 유역면적별 계류차수로 구성되며 출력은 1회 토사량, 관성가중치, 정지조건 의 세 가지로 RWM의 파라미터에 해당된다. 학습 결과는 일반화 오류(generalization error)를 측정하여 평가하는 것이 일반적이며, 이 때 주어진 샘플을 학습용과 테스트용으로 나누어 진행하는 cross-validation을 따른다. 데이터의 개수가 적은 경우에는 테스트 데이터를 하나의 데이터 포인트로 구성하는 leave-one-out 방법을 사용한다. 이보다 데이터가 더 적은 경우에는 동일한 샘플 집합을 학습과 테스트에 동일하게 사용하는 test-on-train 방법을 사용할 수 있다. Test-on-train은 동일한 샘플을 학습과 테스트에 사용하기 때문에 일반화 오류를 측정할 수 없어 객관적인 평가 방법으로는 적합하지 않지만, 너무 적은 데이터로 학습을 진행하는 경우 학습된 모델 자체의 신뢰성이 떨어지므로 다른 방법을 사용할 수 없다. 이 논문에서도 test-on-train 방법을 사용하였다.

데이터는 먼저 정규화를 시행하였다. 신경망에서는 일반적으로 출력이 $[-1, 1]$ 로 제한되므로 입력과 출력의 모든 차원을 독립적으로 식 (3)을 사용하여 정규화하였다.

$$x' = \frac{x - \mu_x}{2\sigma_x} \quad (3)$$

이 때 μ_x 와 σ_x 는 데이터 x 의 평균과 표준편차를 나타낸다. 입력과 출력이 주어진 경우 먼저 은닉층의 개수와 각 은닉층의 노드 개수를 결정하여야 한다. 그림 5는 은닉층과 은닉층의 노드 개수를 변화시켜 가면서 얻어진 평균 오류는 나타낸다. 그림 5에서 x 축은 모델 번호를 나타내며 모델 번호가 증가할수록 은닉층과 노드 개수가 증가한다. 즉, 모델의 복잡도가 증가한다. y 축은 평균 오류로 식 (2)에 따라 계산되었다.

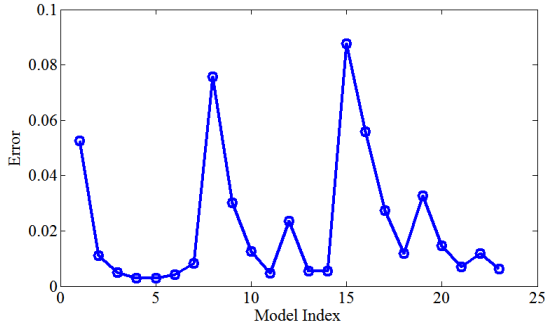


그림 5. 모델에 따른 자승 오류의 평균
Fig. 5. Mean square errors with respect to model index

신경망의 학습은 국부 최적해에만 수렴하며 토석류 데이터의 경우 그 개수가 적기 때문에 잘못된 국부 최적해에 수렴할 수도 있다. 이 실험의 목적은 주어진 데이터를 정확하게 묘사할 수 있는 신경망의 구조를 찾아내기 위함이므로 그림 5의 결과는 자승 오류의 합이 1보다 큰 경우에는 수렴에 실패한 것으로 간주하여 제외하고 100회 반복 실험하여 평균한 값이다. 오류의 합이 1보다 큰 경우 수렴에 실패한 것으로 간주한 것은 반복적인 실험을 통해 오류의 분포를 가우시안 함수로 추정하고 2σ 이하의 오류를 나타내는 경우만을 테스트에 사용하기 위함이다. 복잡한 구조의 신경망의 경우 작은 오류값을 보일 수는 있지만 구조가 복잡해질수록 많은 학습 데이터를 필요로 하고 일반화 특성이 나빠지므로 그림 5에서 오류가 국부적으로 최소가 되는 가장 간단한 구조, 한 개의 은닉층과 5개의 은닉 노드를 가지는 구조를 선택하였다. 선택된 구조는 신경망과 의사 샘플 신경망의 기본 구조로 사용하였다.

토석류 데이터에서 의사 샘플은 데이터 포인트와 데이터 집합의 공분산 행렬을 이용하여 다차원 가우시안 분포로부터 생성하였다. 이는 식 (1)과 동일하지만 토석류 데이터에서는 원 샘플의 중요성을 고려하여 원본 샘플을 N_O 만큼 반복해서 첨가하였다. 결과적으로 토석류 데이터에서 사용한 학습 샘플 X_{PS}' 는 식 (4)와 같이 나타낼 수 있다.

$$X_{PS}' = X_{PS} \cup \left\{ \bigcup_{i=1}^{N_O} X \right\} \quad (4)$$

이 때 데이터 포인트 당 의사 샘플 개수 $N_{PS} = 30$ 으로 설정하였고, 의사 샘플에 의한 편증을 막기 위한 원 샘플의 반복 회수 $N_O = 10$ 으로 설정하였다.

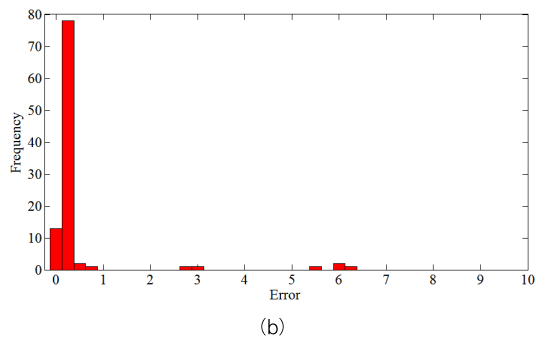
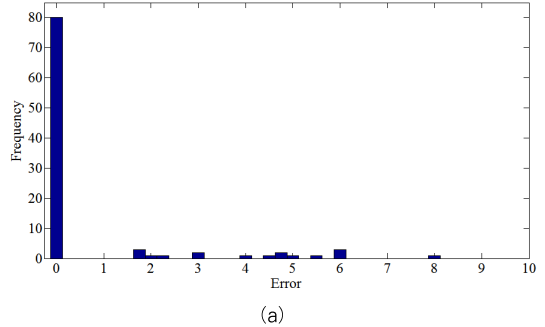


그림 6. (a) 신경망과 (b) 의사 샘플 신경망을 이용한 경우의 오류 히스토그램
Fig. 6. Error histograms using a neural network (a) and a pseudo sample neural network (b)

그림 6은 신경망과 의사 샘플 신경망의 결과를 비교한 것으로 100회 반복 실험한 오류를 히스토그램으로 나타낸 것이다. 그림 6에 나타난 바와 같이 신경망은 오류가 영인 경우가 대부분이다. 즉 주어진 입력과 출력을 정확히 학습하고 있다. 하지만 이는 신경망이 학습 데이터에 지나치게 적응된 (over-fitting) 결과이다. 이 경우 일반화 오류(generalization error)는 크게 나타나는 것이 일반적이다. 이 논문에서는 충분한 샘플이 확보되지 않아 일반화 오류는 비교할 수 없었다. 이에 비해 의사 샘플 신경망은 대부분이 작은 오류 값을 나타내고 있다. 이는 의사 샘플에 의한 평탄화(smoothing) 효과 때문으로 일반화 오류를 감소시키는 역할을 한다. 또한 신경망의 경우 국부 최적해에 빠져 큰 오류 값을 나타내는 경우가 많은 반면 의사 샘플 신경망은 의사 샘플의 평탄화 효과로 국부 최적해에 빠지는 확률이 적음을 알 수 있다. 그림 6의 실험 결과를 요약한 것이 표 2로 의사 샘플 신경망에서 평균 오류와 오류의 분산이 모두 작음을 확인할 수 있다.

표 2. 신경망과 의사 샘플 신경망의 학습 결과 요약
 Table 2. Summary of the training results of neural networks and pseudo sample neural networks

	신경망	의사 샘플 신경망
평균 오류	1.1010	0.4941
오류 분산	7.8466	1.4208

V. 결론 및 향후 연구 과제

이 논문에서는 토석류 퇴적 모델에 사용하는 Random Walk Model(RWM)의 파라미터 값들을 주어진 데이터로부터 자동으로 결정할 수 있는 방법을 제안하였다. 특히 토석류 데이터의 획득이 어려운 점을 고려하여 학습 샘플의 부족을 완화하는 방법으로 의사 샘플을 활용하여 신경망을 학습시키는 의사 샘플 신경망(pseudo sample neural network)을 제안하였다. 의사 샘플은 기존 샘플과 가우시안 분포를 이용하여 생성된다. 이처럼 생성된 의사 샘플은 해공간을 평탄화하고 국부 최적해에 빠질 확률을 줄여줌으로써 보다 안정적인 신경망의 학습이 가능하도록 해주며 이는 RWM의 파라미터 값들을 안정적으로 예측할 수 있도록 해준다. 이러한 사실들은 인공 데이터와 토석류 데이터를 이용한 실험 결과를 통해 확인할 수 있다.

비록 의사 샘플 신경망이 기존 신경망에 비해 학습 샘플의 수가 적은 경우에도 사용할 수 있는 방법이지만, 토석류 데이터의 경우 데이터의 수가 적어 일반화 오류를 비롯한 충분한 검증을 수행할 수 없었다. 현재 추가적인 데이터 획득이 진행 중이며 진행 과정에서 얻어진 데이터를 이용한 실험에서도 이 논문에서와 유사한 결과를 얻을 수 있었다. 향후 추가된 데이터를 이용하며 엄밀한 검증을 진행할 예정이다.

참고문헌

- [1] L. Grady, "Random Walks for Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 11, pp. 1768-1783, Nov. 2006.
- [2] Chang-Woo Lee, Choongshik Woo, and Ho-Joong Youn, "Analysis of Debris Flow Hazard Zone by the Optimal Parameters Extraction of Random Walk Model - Case on Debris Flow Area of Bonghwa County in Gyeongbuk Province," Journal of Korean Forest Society Vol. 100, No. 4, pp. 664-671, Apr. 2011.
- [3] R.P.W. Duin, "Small sample size generalization," Proceedings of the 9th Scandinavian Conference on Image Analysis, pp. 957-964, Oct. 1995.
- [4] S. Haykin, "Neural Networks: A Comprehensive Foundation," 2nd ed. Prentice Hall, 1998.
- [5] C.M. Bishop, "Pattern Recognition and Machine Learning," 2nd ed. Springer, 2007.
- [6] R. Polikar, L. Udpa, S.S. Udpa, and V. Honavar, "Learn++: An Incremental Learning Algorithm for Supervised Neural Networks" IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews, Vol. 31, No. 4, pp. 497-508, Aug. 2001.
- [7] D. Foley, "Considerations of sample and feature size," IEEE Transactions on Information Theory, Vol. 18, No. 5, pp. 618-628, Oct. 1972.
- [8] S. Uchimura, Y. Hamamoto, and S. Tomita, "Effects of the sample size in artificial neural network classifier design," Proceedings of the IEEE International Conference on Neural Networks, pp. 2126-2129, Dec. 1995.
- [9] T.G. Niel, T.R. McVicar, and B. Datt, "On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification," Remote Sensing of Environment, Vol. 98, No. 4, pp. 468-480, Oct. 2005.
- [10] D. Richard, "Probability: Theory and Examples," 4th ed. Cambridge University Press, 2004.

저 자 소 개



허 경 용

1994 : 연세대학교 전자공학과 공학사

1996 : 연세대학교 전자공학과 공학석사

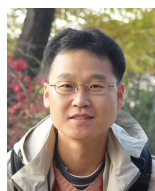
2009 : University of Florida

컴퓨터공학과 공학박사

현 재 : 동의대학교 전자공학과 교수

관심분야 : 인공지능, 패턴인식, 로봇공학

Email : hgycap@deu.ac.kr



이 창 우

1992 : 영남대학교 산림자원학과 농학사

1994 : 영남대학교 산림자원학과 농학석사

1998 : 일본 동경대학

농학생명과학연구과

산림과학전공 농학박사

현 재 : 국립산림과학원

관심분야 : 토석류, 시방, 산사태, 토질역학

Email : leecw@forest.go.kr



박 충 식

1985 : 한양대학교 전자공학과 공학사

1987 : 연세대학교 전자공학과 공학석사

1994 : 연세대학교 전자공학과 공학박사

현 재 : 영동대학교 스마트IT학부 교수

관심분야 : 인공지능, 패턴인식, 인지과학

Email : leciel@yd.ac.kr