

서포트 벡터 머신을 기초로 한 무선 센서 네트워크 환경에서 위치 추정 향상 방안 연구

임재훈, 박귀태  
고려대학교 전자전기공학과

The Research for an enhanced Localization in Wireless Sensor Networks based on Support Vector Machines

Jae-Hoon Lim, Gwi-Tae Park  
Dept. of Electrical Engineering, Korea University

**Abstract** - 현재 센서 네트워크에서 센서의 위치를 추정(Localization)하고자 하는 많은 방법들이 나와 있고, 계속해서 연구 주제로 다루어지고 있다. 이 논문에서는 서포트 벡터 머신(SVM)의 기본적인 내용과 센서 네트워크 분야에서 위치 추정 분야에서 다루어지고 내용들을 서술하고 마지막으로 서포트 벡터 머신을 이용하여, 개선되고 향상된 algorithm을 제시하고자 하는 것이 아닌 SVM을 이용한 적용 사례들과 연구 동향들에 대해 살펴본 뒤 그것들의 적용방법들과 갖는 한계점들, 그리고 그것을 이용한 미래에 연구방향에 대해 고찰해 보고자 한다.

1. 서 론

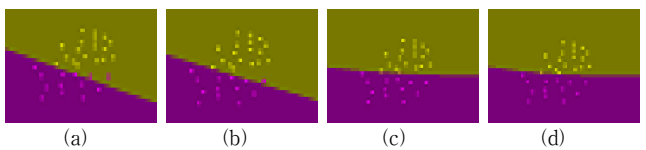
RFID와 USN을 이용한 유비쿼터스 시대의 도래는 생활 속의 곳곳에 이용되어 각각의 기능을 맡은 여러 노드들 끼리 이어져 의사소통이 가능해지고, 이전에는 볼 수 없었던 보다 다양한 서비스의 제공이 가능하게 되었다. 이러한 기능들을 잘 유지하고 지속하기 위해서 센서 노드에 필요한 여러 가지 제약요소들이 따른다. 장애 허용성, 확장성, 전원 관리, 자기 구성 능력, 네트워크 망의 유지, 노드의 생산 가격 등 여러 가지가 그것이다. 그 중에서도 어떤 이벤트가 발생한 지역의 정보를 감지하고 자신의 위치 정보를 이용하여 싱크, 앵커 노드 또는 베이스 스테이션으로 보내도록 하여 그 센싱된 정보의 가치를 보다 높여줄 수 있는 위치 추정의 방법도 중요한 고려 사항 중의 하나이다.

본 논문에서는 특히 위치 추정 방법의 기술에 중점을 두고, 보다 효율적인 위치추정을 위해서 서포트 벡터 머신의 알고리즘을 이용한 사례들을 통해 내용을 전개하고자 한다.

2. 본 론

2.1 서포트 벡터 머신의 분류 방법

신경망(Neural Network)을 이용하여 data를 분류(Classification)하는 기법들은 MLP(Multi Layer Perceptron), K-means Algorithm, EM(Expectation Maximization), SVM 등 여러 가지가 있다. 센서 네트워크에서는 상당한 메모리의 제약이 뒤 따르고 가능한 한 연산을 줄이면서 효율적인 성능을 얻어야 한다. 왜냐하면 센서네트워크에서 사용하는 ISM 밴드에서 다른 대역폭(400MHz, 900MHz) 보다 비교적 전송속도가 높게 나오는 2.4~5 GHz 대역에서 Maximum Bit Rate가 대략 250 Kbps 에 불과하기 때문이다. 또한, 계산 량이 많아지면 자연스럽게 배터리의 빠른 고갈로도 이어지기 때문에 배터리의 수명이 센서 네트워크에서는 치명적인 요소로 작용할 수 있다. 때문에 계산 량이 비교적 적으면서 어느 정도의 성능을 보장할 수 있는 SVM 방법을, 이러한 센서네트워크 분야에 적합하게 적용할 수 있다. 왜냐하면 여러 Training된 Data Set 중에서, 우리는 단지 Hyper Plane에 가장 근접하게 위치하고 있는 Support Vector들을 찾으면 나머지 vector들은 고려를 하지 않아도 되기 때문이다. 예를 들어 1000개의 data가 있어도 Support Vector가 50개 라는 것을 알면 그 나머지 부분에 대해서는 하나하나 계산을 할 필요가 없다.



<그림 1> RBF를 이용하여 Class 값 변경. 차례대로 (a) 5 (b) 10 (c) 100 (d) 200

위의 그림을 통해 LIBSVM [1]을 이용해서 시뮬레이션 해 보았을 때 (c)와 (d)가 거의 차이가 나지 않는 것을 보면 알 수 있듯이, 어느 정도 Class값이 커지면 전체 성능에 큰 차이를 보이지 않는다. 즉, 모든 연산

을 다 해줄 필요가 없다. 또, 다른 분류 방법들과는 달리해야 할 Data가 많아져도 계산 량은 비교적 Linear하게 증가하기 때문에 Exponential 하게 증가하여 소프트웨어적인 성능으로 Cover하기 힘든 다른 알고리즘에 비해 효율적이다.

SVM에서 우리의 목적은 Training된 Data Set의 평균적인 오(誤) 분류 Error를 최소화 시킬 수 있는 Hyper Plane을 찾는 것이다. 이것을 찾기 위해 Lagrange function과 Kuhn-Tucker Theorem을 만족시키는 Parameter들을 찾는 것이다. 이것은 QP(Quadratic programming) Optimization Problem을 통해서 구해진다. 예를 들어 2 차원 상의 Data 로 가정하면 Dual problem을 Maxmization시키는 문제를 푸는 것이 되고 다음과 같은 식으로 나타낼 수 있다.

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j K(x_i, x_j) \quad (1)$$

$$\text{s.t. } w^* = \sum_{i=1}^N \alpha_i^* d_i x_i \quad (2)$$

$$\sum_{i=1}^N \alpha_i^* d_i = 0 \quad (3)$$

$$\alpha_i \geq 0, \text{ and } i = 1, 2, \dots, N \quad (4)$$

여기서  $x_i$  는 Input pattern을 나타내고,  $y_j$  는 Target output을,  $i, j$  는 1부터 N차원까지 하나씩 증가함을 나타낸다. 여기서  $K(x_i, x_j)$  는 라그랑지 함수를 풀면서 얻어지는 제약 상수이다.  $K(x_i, x_j)$  는 Kernel Function을 나타낸 것이다. SVM에서 많이 쓰이는 커널 함수에는 크게 RBF(Radial Base Function), MLP, Polynomial 이 있다. 이들은 Mercer's Theorem을 만족하는 것들로써, Input Space에서 Non-Linear 특징을 가지는 Feature Space로 Mapping 할 때 Inner Product의 특성을 만족하는 것들로 이루어져 있다. 이런 기본적인 것들을 바탕으로 노드들을 분류하고 위치를 추정하는데 쓰이고자 하는 것이다.

2.2 센서 네트워크에서 위치 추정(Localization)의 여러 가지 방법

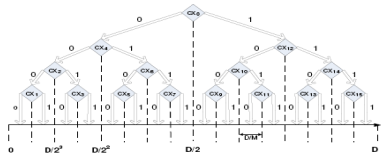
Localization이란 노드의 위치 정보를 알기 위해 구연되어지는 알고리즘 또는 그것을 찾기 위한 과정을 포함하는 개념이라 할 수 있다. 필요한 곳에 노드들이 배치되고, 그것이 센싱한 정보를 모니터링을 함으로써 사용자에게 신뢰하고 정보성이 있는 Data를 주고자 하기 위함이다. 또한 이벤트 탐지(Event detection), 위치기반 계산(Location-dependent computing), 지리적 라우팅(Geographic routing), 커버리지 트래킹(Coverage and tracking) 등 효율적인 네트워크 구축을 위해서도 사용된다[2]. 센서 네트워크에서 사용하는 Localization 방법에는 여러 세부적인 분류방법들이 있지만, 일반적으로 크게 2가지로 분류되어 지고 있다. Range-Based 와 Range-Free 방법이다. Range-Based의 방식은 거리, 각 등과 같은 물리적인 정보를 측정하여 위치를 추정하는 방법이다. 때문에 하드웨어 적으로 다루어지고 상대적으로 대량의 노드들이 필요한 센서 네트워크에서는 Cost관점에서는 그다지 효율적이지 못한 방법이다. Range-Based에 대표적인 것들은 TOA(Time of Arrival)[3], TDOA (Time Difference Of Arrival)[4], AOA(Angle of Arrival)[5], RSSI (Received Signal Strength Indication)[6] 등과 같은 것들이 있다. TOA 방법은 전파의 도달 시간을 이용하여 거리를 계산하는 방법이다. TDOA는 신호의 도착시간 차이를 이용하는 방법이며, AOA는 신호가 들어오는 곳의 각도 정보를 가지고 위치를 예상하는 방법이다. RSSI는 수신하는 쪽에서 신호의 세기를 측정하여 확률적인 방법으로 구축되었던 표본들과 비교하여 위치를 측정한다.

Range-Free방법에서 대표적인 것들은 Centroid Algorithm[7], DV-Hop Algorithm[8], APIT[9] 등이 있다. 이는 상대적으로 센서 필드에 무작위

로 배치되는 다수의 센서노드에 적용할 때 유리한 점을 가지고 있다. 하지만 하드웨어를 이용하지 않은 소프트웨어적으로 구현되기 때문에 정확한 위치를 측정하는데 있어서는 아무래도 Range-Based보다 부정확하다 할 수 있다. Centroid는 미리 앵커 노드들이 주어진 상태에서 더 이상 새로운 Data가 들어오지 않는다고 판단하면 앵커 노드들과 그 개수를 나눈 값으로 Estimation하게 된다. DV-Hop은 센서 노드들과 앵커 노드 사이의 Hop을 Count함으로써, 한 Hop의 평균 거리를 통하여 삼각측량(Tri-Literation)으로 위치를 구하는 방식이다. APIT에서는 모든 앵커 노드들이 자기의 위치정보를 Broadcast하고, 노드들은 자신이 가지고 있는 앵커 테이블을 옆 노드들과 교환한다. 그리고 Random 임의의 3개 앵커로부터 가상의 선을 이어 자신이 그 삼각형 안에 들어 있는지 판단하고 겹치는 곳이 생기는 부분을 찾음으로써 위치를 결정하게 된다. 위에서 언급된 것들은 이미 널리 알려진 것들이고, 이 외에도 이동성을 가진 비콘 노드를 이용한 Localization 기법[10], 이동성 로봇을 이용한 센서 위치 인식 기법 연구[11] 등 Localization 관련 연구들이 활발히 진행되고 있다.

### 2.3 소프트 벡터 머신을 기초로 한 위치 추정 분야의 적용

앞에서 살펴본 바와 같이 파워와 가격의 제약이 없다면 모든 노드에 특별한 하드웨어(예를 들어 GPS)를 부착하여 위치를 측정하면 좋겠지만, 그렇지 못한 것이 현실이다. 그래서 상대적으로 경제적인 Range Free 방법을 이용한 알고리즘이 많이 연구되고 있고, 그중에서 Localization을 하는데 SVM을 적용한 논문이 제안되어 있다[12]. SVM에 대한 Training Set을 Beacon node들의 Set으로 잡고, Kernel Function은 노드의 Hop Count로 설정하고 있다. 비콘(Beacon) 노드는 절대적인 위치를 알 수 있기 때문에 그 수가 많으면 많을수록 정확도는 높아지게 되고, 노드와 노드사이의 경계에서 sensing이 되지 않아 발생하는 문제인 경계 문제(border problem)에도 효과적으로 대처할 수 있다. 이 논문의 기본 개념은 우리가 알고자 하는 노드의 위치를 아래와 같은 Decision Tree를 이용하여 위치를 찾는다.

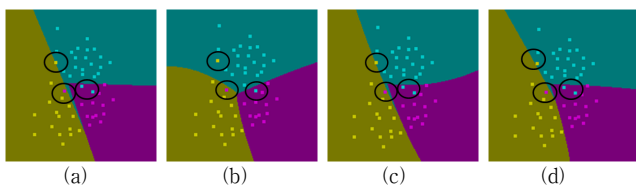


〈그림 2〉 m=4일 때 Decision Tree

위 그림은 비콘이 아닌 노드를 분류하기 위해  $2^m-1$ 개의 Class Set을 고려하고 있다. Parameter m은 우리가 얼마나 센서에 가깝게 위치시킬 수 있는가를 조절하는 요소이다. 위치를 찾는 과정은 크게 3단계로 나누고 있다. 첫 번째로 Training 단계로 주어진 비콘 노드 Set을 이용하여 Training을 시키고, 두 번째는 Advertisement 단계로 Broadcasting을 이용하여 비콘에서 각 노드들로 위치 정보를 전파한다. 마지막은 Localization 단계로 2단계를 통하여 얻어진 위치 정보를 바탕으로 이미 Training된 Data Set에 적용하는 것이다. 여기서 고려된 Error는

$$- E^u = \sqrt{2}D \left( \frac{1}{2^m} + \frac{7}{8} - \frac{(1-\epsilon)^m}{2^{m+1}} - \frac{(2-\epsilon)^m}{2^m} + \frac{(4-3\epsilon)^m}{2^{2m+3}} \right) \quad (5)$$

이다. D는 Tree의 길이,  $\epsilon$ 은 원래 SVM이 가지고 있는 에러, 그리고 Parameter m을 조절하여 원래 노드의 정확한 위치와 추정된 위치의 에러를 Modified Mass-Spring Optimization을 이용하여 최소화 한다. 그렇게 함으로써 노드의 위치를 찾겠다는 것이다. 그런데 이 논문에서는 몇 가지 한계점들을 찾아 볼 수가 있다. 먼저, 이 논문에서는 3가지 가정을 전제로 하고 있는데 첫째로 비콘 노드가 있어야 하고, 둘째로 센서는 어떤 비콘 노드로부터 바로 정보를 받지 못할 수도 있다는 점, 셋째로 단지 연결정보는 위치추정을 해야 하는데만 사용될 수 있다는 점인데, 실제로 센서 필드에 대규모로 배치되는 특성을 고려해볼 때 상당히 제약적인 조건이 아닐 수 없다. 또 이 논문에서는 Kernel Function으로 RBF만을 사용하여 작성되었다. 다음은 시뮬레이션을 통해 각각의 Kernel Function들에 대해 비교해 본 그림이다.



〈그림 3〉 Kernel Function 들에 대한 비교

Mercer's Theorem을 만족하는 Kernel Function들 중에 (a) Linear (b) Polynomial (c) Radial Basis Function (d) Sigmoid를 각각 이용하였고 class는 3, 노드 수는 1000개를 사용하였다. 여기서 검정색 동그라미는 노드의 Misclassification을 나타낸 것이고, (b)의 경우(Polynomial)의 경우가 RBF보다 에러가 적은 것을 확인할 수가 있다. 대부분의 경험적인 방법에 의해서 RBF를 많이 사용하고 있지만 꼭 이것이 최선의 선택이라고는 말할 수 없고, 이 논문에서는 다른 커널 함수를 사용하지 않아 비교 대상이 없으므로 그 결과가 좋은지 나쁜지 상대적으로 비교하기 힘들다. 그리고 이 논문의 가정에서 비콘 노드가 처음에는 존재한다 하더라도 그들의 센서 네트워크의 특성 상 짧은 배터리 수명으로 인해 그 노드가 죽으면 구성된 Training Data의 재수정을 해야 하는 번거로움이 생기는 등의 한계점 등이 있다.

이 논문의 기본 이론을 적용하여 Mobile Nodes에 대해 새로운 Localization 알고리즘을 적용한 논문도 있다[13]. 이 논문은 RSSI를 이용한 RF-방법 단계, 그리고 연결 정보와 SVM 알고리즘을 가지고 Area Localization하는데 적용함으로써 궁극적으로 Localization Error를 줄이고자 노력하였지만 역시 실제 적용의 한계점이 존재한다.

### 3. 결 론

본 논문에서는 SVM 알고리즘을 이용하여 센서 네트워크의 Localization 분야에 적용한 논문들을 고찰해 보았다. Range Based 방법과 비교하여 볼 때, 소프트웨어 적인 방법으로서 최선의 알고리즘을 구현하여 Localization하는 Range Free방법에서는 최대한의 위치 에러를 줄이는 것이 목적이다. 위의 SVM을 적용한 논문들에서는 에러를 줄이는 효과를 가져왔으나 극히 제한적인 조건에서만 가능한 실험 결과들이고 그것을 센서 필드에 까지 적용하기에는 아직 많은 한계점들이 있어 보인다. 아직까지는 센서 네트워크 분야에서 SVM이론을 적용할 수 있는 영역은 SVM이 가지는 특성 때문에 이 분야의 여러 방면에 적용하기는 쉽지 않다. 그렇기 때문에 SVM 방법만 가지고 해결책을 찾기 보다는, 그 이외에도 다른 여러 분야와의 이론들을 접목하여 추후 연구가 이루어져야 더욱 더 좋은 결과를 낼 수 있을 것이라 생각된다.

### 참 고 문 헌

[1] Chang, C., Lin, C. "LIBSVM -- A library for Support Vector Machines", National Taiwan University, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

[2] 최상호, "유비쿼터스 센서 네트워크의 위치탐지 기술 및 동향", 주간기술동향 통권, 1291호, P27, 2007.04.11

[3] Harter, A.et al. "The anatomy of a Context-aware application", Proceedings of MOBICOM 1999, Seattle, Washington, pp. 59-68 ,1999

[4] Girod, L.et al. "Robust range estimation using acoustic and multimodal sensing." ,IROSO 2001. Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems, vol.3, pp. 1312-1320. IEEE Robotics and Automation Society, Maui (2001)

[5] Niculescu, D.et al. "Ad hoc positioning system (APS) using AoA." ,Proceedings of IEEE INFOCOM 2003, pp. 1734-1743 (2003)

[6] Girod, L. et al. "Locating tiny sensors in time and space: A case study." ,Werner, B. (ed.) Proceeding of 2002 IEEE International Conference on Comoyter Design: VLSI in Computers and Processors, pp. 214-219. IEEE Computer Society, Freiburg, Germany (2002)

[7] N. Bulusu et al. "GPS-less low cost outdoor localization for very small devices" IEEE Personal Communications Magazine, vol7, no.1,pp267-280,2003

[8] Niculescu, D.et al."Dv based positioning in ad hoc networks." J. Journal of Telecommunication Systems 22, 267-280 (2003)

[9] He, T. et al. "Range-free localization schemes in large scale sensor networks.", Proceedings of the Ninth Annual International Conference on Mobile Computing and Networking, pp. 81-95 (2003)

[10] 이병화, "무선 센서 네트워크에서 거리 정보를 필요로 하지 않는 위치 인식 알고리즘", 석사학위논문, 고려대 대학원 전자전기공학과, P1-50, 2007.12

[11] 김우현, "무선 센서 네트워크에서 이동성 로봇을 이용한 센서 위치 인식 기법에 관한 연구", 대한생산응용학회논문집, 제2호, P105-113, 2007.05

[12] Duc, A., Nguyen, T."Localization in Wireless Sensor Networks based on Support Vector Machines", Submitted to IEEE Transactions on Parallel and Distributed Systems, (2007)

[13] H.Zhang et al. "Area Localization Algorithm for Mobile Nodes in Wireless Sensor Networks Based on Support Vector Machines," MSN2007, LNCS 4864, P561-571, 2007