

# FCM 클러스터링과 다변량 퍼지결정트리를 이용한 상황인식 보안 서비스

양석환\*, 정목동\*\*

\*부경대학교 정보보호학 협력과정

\*\*부경대학교 컴퓨터공학과

e-mail:tigergal@chol.com, mdchung@pknu.ac.kr

## Context-Aware Security Service using FCM Clustering and Multivariate Fuzzy Decision Tree

Seokhwan Yang\*, Mokdong Chung\*\*

\*Dept of Information Security, Pukyong National University

\*\*Dept of Computer Engineering, Pukyong National University

### 요 약

유비쿼터스 환경의 확산에 따른 다양한 보안문제의 발생은 센서의 정보를 이용한 상황인식 보안 서비스의 필요성을 증대시키고 있다. 본 논문에서는 FCM (Fuzzy C-Means) 클러스터링과 다변량 퍼지 결정트리 (Multivariate Fuzzy Decision Tree)를 이용하여 센서의 정보를 분류함으로써 사용자의 상황을 인식하고, 사용자가 처한 상황에 따라 다양한 수준의 보안기술을 유연하게 적용할 수 있는 상황인식 보안 서비스를 제안한다. 제안 모델은 기존에 많이 연구되어 오던 고정된 규칙을 기반으로 하는 RBAC(Role-Based Access Control)계열의 모델보다 더욱 유연하고 적합한 결과를 보여주고 있다.

### 1. 서론

최근 확산되고 있는 유비쿼터스 환경 구축은 개인정보 유출 및 다양한 보안문제를 발생시켰고, 이에 따라 유비쿼터스 환경에서의 보안기술 연구가 점차 주목받고 있다. 유비쿼터스 환경 구축의 기반을 이루는 요소 중의 하나인 센서 네트워크는 온도, 습도, 조도, 영상 및 사용자의 이동 상황, 건강상태와 같은 정보를 시스템에 제공하는 다양한 센서들로 구성되어 있으며, 이러한 센서들이 시스템에 제공하는 사용자의 주위환경 및 상황에 대한 정보는 사용자의 프라이버시 침해 문제를 일으킬 가능성이 크다. 따라서 이러한 정보의 접근에 대하여 강력한 보안기능을 제공할 필요가 있다.

그러나 모든 상황에서 강력한 보안기술을 적용하는 것은 오히려 시스템의 효율성 저하와 함께 사용자의 불편을 가중시킬 수 있으며, 특히 응급구조 시스템과 같이 긴급 상황 시에는 보안보다 사용자의 생명을 보호하는 일을 우선적으로 고려해야 하는 시스템도 존재한다. 이에 본 논문에서는 FCM (Fuzzy C-Means) 클러스터링[1] 과 다변량 퍼지결정트리 (Multivariate Fuzzy Decision Tree : MFDT)[2]를 이용하여 센서의 정보를 분류함으로써 사용자의 상황을 인식하고, 사용자가 처한 상황에 따라 다양한 수준의 보안기술을 유연하게 적용할 수 있는 상황인식 보안 서비스를 제안한다.

본 논문의 구성은 다음과 같다. 2절에서는 관련 연구에 대하여 살펴보고, 3절에서는 FCM 클러스터링과 다변량 퍼지결정트리를 이용한 상황인식 보안 서비스 모델을

소개한다. 4절에서는 제안한 상황인식 보안 서비스 모델을 직접 구현하고 실험한 결과 분석을 다루고 마지막으로 5절에서는 결론 및 향후 연구 방향을 제시한다.

### 2. 관련연구

상황인식에 관한 연구는 마크 와이저의 유비쿼터스 컴퓨팅 개념이 제시된 이후로 많은 연구가 진행되어 왔다. 상황인식 기술을 보안에 적용한 대표적인 예로서 Illinois 대학에서 수행하고 있는 Gaia[3] 프로젝트의 Cerberus와 Georgia Tech에서 제안한 CASA (Context-Aware Security Architecture)[4] 등을 들 수 있다.

Cerberus는 ACL (Access Control List)을 기반으로 접근제어를 수행한다. ACL은 전통적인 접근제어 정책 모델로서 간단한 정책의 구성이 가능하지만 데이터 추상화 기능을 제공하지 않기 때문에 보안 서비스를 제공하고자 하는 조직의 권한 및 책임 구조를 보안 정책에 적용하기 어렵다.

CASA는 GRBAC (Generalized Role-Based Access Control)[5,6]모델을 기반으로 접근제어를 수행하므로 사용자, 객체, 환경정보들을 역할로 추상화하여 보안 정책 관리기능을 제공한다. 그러나 GRBAC는 RBAC[7]가 제공하는 권한 추상화를 제공하지 않으며, 관리 도메인 요소들이 많아짐에 따라서 관리해야할 역할의 양이 많아진다. 또한 객체 역할과 환경 역할 간의 중복되는 영역에 의해 발생하는 충돌 문제로 인해 보안정책 관리 및 구성에 따르는 오버헤드가 많이 발생하게 되며, 보안정책의 설정 여부에

따라 예상하지 못한 상황, 즉 보안 정책에 등록되지 않은 상황에 대한 대응이 어렵다는 단점을 가진다.

현재 수행되고 있는 상황인식 보안연구는 Cerberus 및 CASA와 같이 ACL, GRBAC 등을 적용하고 있는 경우가 많으며, Cerberus 및 CASA와 동일한 문제에 노출되어 있다. 따라서 이러한 문제점을 해결할 수 있는 상황인식 보안 서비스에 대한 연구가 요구된다.

### 3. FCM 클러스터링과 다변량 퍼지결정트리를 이용한 상황인식 보안 서비스 모델

본 논문에서는 FCM 클러스터링과 다변량 퍼지결정트리를 이용하여 입력된 데이터에 대한 패턴 분류를 수행하고 그 결과를 이용하여 보안 등급을 결정함으로써 다양한 환경에 대하여 유연한 보안 서비스를 제공하는 상황인식 보안 서비스 모델을 제안한다.

#### 3.1 FCM 클러스터링

FCM (Fuzzy C-Means)[1]은 하나의 클러스터에 속해 있는 각각의 데이터 점을 소속정도에 의해서 분류하는 데이터 분류 알고리즘이다. FCM 클러스터링은 퍼지 분할 기법을 사용하며 소속 함수 U는 0과 1사이의 값을 가지는 요소들을 가진다. 데이터 집합에 대한 소속정도 값의 합은 항상 1이다.

$$\sum_{i=1}^c u_{ik} = 1, \forall k = 1, \dots, n \quad 0 < \sum_{k=1}^n u_{ik} < n$$

FCM 클러스터링에 대한 비용함수는 다음과 같은 형태를 가진다[8].

$$J(u_{ik}, v_i) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m (d_{ik})^2$$

$$v_i = \{v_{i1}, v_{i2}, \dots, v_{ij}, \dots, v_{iL}\}$$

$$d_{ik} = d(x_k - v_i) = \left[ \sum_{j=1}^L (x_{kj} - v_{ij})^2 \right]^{\frac{1}{2}}$$

$$v_{ij} = \frac{\sum_{k=1}^n (u_{ik})^m x_{kj}}{\sum_{k=1}^n (u_{ik})^m}, u_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}}$$

$u_{ik}$ :  $i$ 번째 클러스터에 속한  $x_k$ 의  $k$ 번째 데이터의 소속도

$v_i$ :  $i$ 번째 클러스터의 중심벡터

$m$  ( $1 \leq m < \infty$ ): 분류 공정에서 퍼지성의 양을 제어하는 파라미터. 일반적으로  $m = 2$

#### 3.2 다변량 퍼지결정트리[9]

결정트리는 분류규칙을 표현하는 트리이다. 비 단말모드에는 분류를 위해 비교하는 데이터의 특징이 명시되고, 링크에는 비교조건 또는 특징 값이 부여되며, 단말노드에는 루트노드에서 해당 노드까지의 경로 상에 있는 모든

조건을 만족하는 데이터가 속하는 클래스 값이 부여된다. 결정트리는 명확한 값을 기준으로 특징공간을 분할하므로, 미세한 차이를 가지는 서로 다른 두 데이터를 각각 다른 클래스로 분류할 수 있다. 따라서 결정트리에 퍼지함수의 개념을 도입하여, 특징공간을 소속함수를 이용하여 정의한 퍼지 경계면으로 분할하는 퍼지 결정트리에 대한 연구가 진행되고 있다[10].

<표 1> 다변량 퍼지결정트리를 이용한 분류과정[9]

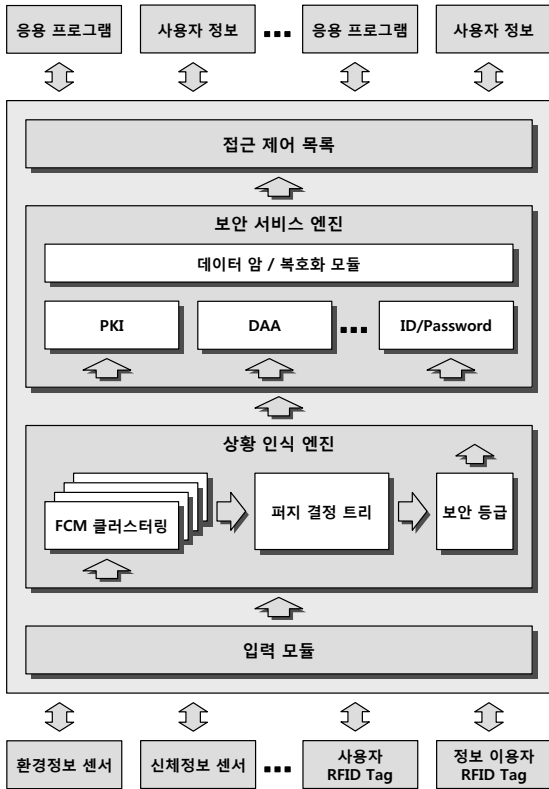
<p>[단계 1] 근 노드에서 각 단말 노드까지 거쳐 가는 노드에서의 소속 함수와 속성 벡터를 이용하여 T-norm을 계산</p> <p>1) 각 노드의 속성 벡터 <math>w_i</math>의 방향으로 입력데이터 <math>x</math>를 사영시킨다.</p> $z_i = w_i^T x$ <p>2) 근 노드에서 <math>n</math>번째 단말 노드까지의 T-norm을 계산</p> $T_n = \prod_{i=\sqrt{\text{node}}}^{n_{th} \leq \text{af node}} \mu_{S_{nw}}^i(Z_i)$ <p>위 과정을 모든 단말 노드에 대하여 반복하면 각 단말 노드는 0에서 1사이의 T-norm을 가진.</p> <p>[단계 2] 단말 노드 중에 각 클래스에 해당하는 단말 노드들의 T-norm의 평균을 계산하고 이 평균 값이 가장 큰 클래스로 입력 데이터를 분류한다.</p> $\text{class } C = \arg_i \max A_i$ <p>(<math>A_i</math>: <math>i</math>번째 클래스를 갖는 단말 노드들의 T-norm의 평균)</p>
---

#### 3.3 제안 모델

본 논문에서는 FCM 클러스터링 알고리즘과 다변량 퍼지결정트리를 적용한 상황인식 보안 서비스 모델을 제안한다.

FCM 클러스터링 알고리즘을 이용한 상황인식 보안 서비스 모델은 분류된 각 클러스터의 중심과의 거리를 이용하여 가장 가까운 분류결과를 적용하므로 예기치 못한 상황에서도 가장 적절한 결과를 도출할 수 있다. 또한 최종적으로 도출된 보안 등급에 대한 접근제어목록을 적용할 경우 다양한 상황의 충돌에 대한 무결성 보장문제도 해결가능하다.

본 논문에서 제안하는 모델은 상황정보의 수집 및 전송을 수행하는 입력모듈, 입력된 정보를 이용하여 상황정보를 분류하고 최종적인 결과를 도출하는 상황인식 엔진, 다양한 보안 기능을 제공하는 보안 서비스 엔진, 해당하는 보안 등급에 따른 접근권한 목록을 관리하는 접근 제어 목록으로 구성된다. 그림 1은 제안 모델의 구조를 나타낸다.



(그림 1) 제안 모델의 아키텍처

4. 구현 및 평가

본 논문에서 제안한 모델의 학습 및 성능을 파악하기 위하여 P-IV 3.0GHz CPU와 1GB Ram이 장착된 PC에서 Java 5.0을 이용하여 구현하였다. 상황 정보의 입력에는 TynyOS 2.x를 기반으로 하는 UBee430/UBee430-AP-Kit을 사용하였다. 실험 데이터는 고혈압 환자 혈압변화에 대한 300건의 가상 데이터를 사용하여 클러스터링을 수행하였고, 다변량 퍼지결정트리를 통하여 어떤 보안 등급을 계산해 내는지 확인하였다. 또한 GRBAC 모델과 FCM 클러스터링 모델을 각각 적용했을 때의 보안등급 계산결과를 비교하였다. 표 2는 GRBAC 모델에 적용하기 위한 권한정책의 예를 나타낸다.

<표 2> GRBAC 모델에서의 권한정책의 예

환경정보		작업	행동 주체	대상	혈압	등급
시간	장소					
모든 시간	모든 장소	정보 조회	의료 요원	환자 정보	156/100~	1
					155/99~151/98	2
					150/97~146/96	3
					145/95~141/94	4
					~140/93	5

보안등급은 총 5등급으로 분류되며 등급이 낮을수록 낮은 수준의 보안 서비스를 적용한다. 표 3은 GRBAC 모델을 적용한 보안등급 계산결과와 FCM 클러스터링 알고리즘을 적용한 보안등급 계산결과를 나타낸다. 표 3에서 기대등급은 정상혈압수치와 비교하여 위험수치로 다가가는 정도에 따라 결정한 등급이다.

<표 3> 혈압변화에 따른 보안등급 계산 결과

혈압	GRBAC	FCM	기대등급
152.5 / 98.8	2	3	3
153.2 / 100.3	Unknown	3	3
167.7 / 109.5	1	1	1
157.3 / 105.3	1	2	2
141.5 / 87.4	Unknown	3	4
137.3 / 85.3	5	4	4

표 3에서 혈압수치 167.7 / 109.5의 경우, 위급 상황으로 인식하여 1등급의 보안서비스가 적용되고 있다. 또한 다소 높은 혈압범위에서는 2~3등급, 정상 혈압 범위에서는 4~5등급의 보안등급을 제시하고 있음을 확인할 수 있다.

그러나 GRBAC 모델을 적용한 보안등급 계산결과에서는 Unknown이 발생하고 있으며, 이는 환자의 혈압상태가 정규적으로 분류되지 않고 두 개 이상의 조건에 교차되어 있는 경우이다. GRBAC 모델의 권한정책을 더욱 상세하게 설정할 경우 적절한 보안등급의 도출이 가능하겠지만 모든 상황에 대하여 적절한 각각의 권한정책 설정은 현실적으로 매우 어려우며 또한 권한정책의 분류 개수가 많아 질수록 결과 도출에 시간이 걸릴 것을 예측할 수 있다.

반면, FCM 클러스터링을 적용한 결과는 일반적으로 모든 상황에 대하여 적절한 결과를 보여주고 있다. 이는 자율적인 클러스터링을 수행한 후, 각 클러스터의 중심 값과 입력벡터의 거리를 이용하여 더욱 가까운 클러스터, 즉 더욱 적절한 등급을 제시하기 때문이며 GRBAC 모델보다 더욱 유연하고 적합한 결과를 보여주고 있다.

또한 제안 모델은 다양한 상황정보를 기반으로 상황인식을 수행하므로 혈압정보 외에도 심박수, 체온, 주변기온 및 습도 등 많은 센싱 정보가 FCM 클러스터링을 통해서 각각의 보안등급으로 분류된다. 따라서 계산된 각각의 보안등급을 통합하여 분석할 필요가 있다. 이에 FCM 클러스터링을 통해 도출된 결과 데이터를 다시 다변량 퍼지결정트리에 적용함으로써 다양한 정보를 통합하여 가장 적절한 보안등급을 선택할 수 있다.

고혈압 환자의 경우 혈압정보 외에 심박수, 체온, 주변 온도 등의 영향을 많이 받으며 정보별로 그 중요도가 다르기 때문에 FCM 클러스터링의 결과에 중요도 값을 적용하여 다변량 퍼지결정트리의 입력 값으로 사용한다. 표 4는 FCM 클러스터링과 MFDT를 이용한 보안등급을 나타낸다.

<표 4> FCM 클러스터링과 MFDT를 이용한 보안등급

혈압	심박수		FDT
	FCM	횟수	
167.7 / 109.5	1	122	1
157.3 / 105.3	2	78	2
120.8 / 82.5	4	73	4
141.5 / 87.4	3	160	3
137.3 / 85.3	4	201	2

표 4에서 혈압수치 137.3 / 85.3의 경우, 혈압 변화에 따른 보안등급은 4등급이지만 심박 수의 변화에 따른 보안등급은 1등급으로 계산되어 최종 보안 등급은 2등급이 제시되고 있음을 확인할 수 있다. 이는 환자의 현재 혈압 상태는 정상이나 높은 심박 수로 인하여 혈압상태가 악화될 가능성이 보안등급에 적용되었음을 나타낸다. 즉 제안 모델은 환자의 주위 상황을 인식함으로써 환자의 상태변화에 대한 파악과 예측을 통해 가장 적절한 보안등급을 제시하고 있음을 확인할 수 있다. 표 5는 기존의 GRBAC 모델과 제안 모델의 비교내용을 보여준다.

<표 5> GRBAC 모델과 제안 모델의 비교

기준	GRBAC 모델	제안모델 (FCM + MFDT)
자율성	사전 등록된 고정규칙기반 상황인식	자율적인 클러스터링을 이용하여 상황인식
대응능력	고정된 규칙을 이용하므로 모든 상황에 대한 대응 불가능	클러스터 중심과의 거리로 분류/인식하므로 모든 상황에 대응가능
관리정책	규칙/정책의 직접 관리로 인한 오버헤드	규칙/정책을 직접 관리할 필요가 없음
통합성	상황정보의 다양성에 의해 정책구조가 갈수록 복잡해 짐	다양한 종류의 상황정보를 통합하여 최적의 결과를 도출

**5. 결론 및 향후계획**

본 논문에서는 사용자의 상황을 인식하고, 다양한 수준의 보안기술을 적용하기 위하여 FCM 클러스터링 알고리즘과 다변량 퍼지 결정트리 이용한 상황인식 보안 서비스를 제안하였다. 제안 모델은 FCM 클러스터링 알고리즘을 통한 자율적인 상황의 분류와 다양한 센서 정보에 대한 분류 결과를 통합하는 다변량 퍼지결정트리를 이용하여 가장 적절한 보안 등급을 제시한다. 그리고 지속적인 관리가 필요한 고혈압 환자에 대한 가상적인 주변상황을 구성하여, 제안 모델이 얼마나 유연하고 적절하게 보안서비스를 적용할 수 있는가에 대한 실험 결과를 제공하였다.

기존의 GRBAC 모델과 비교할 때, FCM 클러스터링을 적용한 모델은 자율적인 클러스터링 방식을 이용하므로 예상하지 못한 상황에 대해서도 적절한 분류결과를 보여주고 있다. 또한 제안 모델에서는 수많은 규칙을 직접 관리하지 않고 자율적으로 상황을 분류하므로 GRBAC 모델과 같은 관리상의 오버헤드 문제도 해결할 수 있었다. 이는 제안 모델이 유연하면서도 자율적인 보안 서비스를 제공함과 동시에 관리에 대한 비용 효율적인 면에서도 좋은 결과를 제공한다는 것을 보여주고 있다.

또한 GRBAC 모델은 상황정보의 종류가 늘어날수록 정책 및 규칙의 구조가 점점 복잡해지지만 제안 모델은 다변량 퍼지결정트리를 이용한 통합결과를 제시함으로써 환자의 상태변화에 따라 예측 가능한 가장 적절한 보안등급을 제시하고 있으며 각 상황정보의 중요도에 따른 등급

결정으로 더욱 현실적인 보안 등급의 적용이 가능하다.

향후 연구는 학습기능을 적용한 보다 현실적이고 사용자 별로 차별화된 성능을 제공하는 시스템으로 발전시켜 나갈 것이며, 보안만이 아닌 상황인식을 필요로 하는 다른 모든 시스템에도 적용 가능한 범용적인 시스템으로 발전시켜 나가야 할 것이다.

**6. Acknowledgements**

본 연구과제는 지역혁신을 위한 대학원생의 양성을 위하여 산업자원부 장관과 2007년도 브레인 부산 21 사업의 지원으로 수행되었음.

**참고문헌**

- [1] J. Bezdek, "A convergence theorem for the fuzzy ISODATA clustering algorithm," *IEEE Trans. Pattern Anal. Machine Intelligence*, Vol.PAMI-2, No.1, 1980, pp.1-8.
- [2] 전문진, 도준형, 이상완, 박광현, 변증남, "다변량 퍼지의사결정트리와 사용자 적응을 이용한 손동작 인식," *로봇공학회 논문지*, 제3권, 제2호, 2008, pp.81-90.
- [3] J.A. Muhatadi, et al., "Cerberus: A Context-Aware Security Scheme for Smart Spaces," In Proc. of the First International Conference on Pervasive Computing and Communications (PerCom'03), 2002.
- [4] M.J.Covington, et al., "A Context-Aware Security Architecture for Emerging Applications," In Proc. of the 18th Annual Computer Security Applications Conferences (ACSAC'02), 2002, pp. 249-258.
- [5] M. J. Convington, et al., "Generalized Role-Based Access Control for Securing Future Applications," In Proc of the 23th National Information Systems Security Conference(NISSC), Baltimore, 2000, pp.115-125.
- [6] M. J. Moyer and M. Ahamad, "Generalized Role-Based Access Control," In Proc of IEEE International Conference on Distributed Computing Systems(ICDSC2001), 2001, pp.391-398.
- [7] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E.Youman., "Role-based access control models," *IEEE Computer*, Vol. 29, No. 2, February 1996, pp. 38-47.
- [8] 오성권, 프로그래밍에 의한 컴퓨터지능, 내하출판사, 2002.
- [9] 전문진, 도준형, 이상완, 박광현, 변증남, "다변량 퍼지의사결정트리와 사용자 적응을 이용한 손동작 인식," *로봇공학회 논문지*, 제3권, 제2호, 2008, pp.81-90.
- [10] 이우향, 이건명, "특징공간을 사전 분할하는 퍼지 결정트리 유도," *정보과학회논문지 : 소프트웨어 및 응용*, 제29권, 제3호, 2002, pp.156-166.