

네이브 베이즈를 이용한 멀티 에이전트의 효율적인 관리 방법 An Effective Management Method of Multi-Agent Using Naive Bayes

황정식, 류경현, 정환목
대구가톨릭대학교 컴퓨터공학부
Jeong-Sik Hwang, Kyung-Hyun Ryu, Hwan-Mook Chung
Faculty of Computer and Information Communication Engineering
Catholic University of Daegu
E-mail: icsman@cu.ac.kr

요 약

멀티 에이전트(Multi-Agent)들이 상호 연동하여 공통의 목적을 수행하기 위해서는 에이전트를 관리하는 매니지먼트 에이전트(Management Agent)가 요구되고, 주어진 환경에서 획득한 제한된 지식을 효율적으로 이용하는 방법이 필요하다.

본 논문에서는 네이브 베이즈 이론을 적용하여 각 에이전트의 속성값(Attribute Value)에 따라 매니지먼트 에이전트가 각 에이전트를 효율적으로 관리할 수 있는 NBMA(Naive Bayes Management Agent)를 제안하고 이를 이용한 미팅 참가 결정 에이전트를 제안한다. NBMA는 고유한 속성을 지닌 여러 개의 하위 에이전트와 그들을 관리하는 매니지먼트 에이전트로 구성되어 있으며 매니지먼트 에이전트는 하위 에이전트들의 고유한 속성에 대한 메타지식을 이용하여 관리 하도록 한다. 하위 에이전트간에는 상호 조건부 독립(mutually conditional independence) 가정하에 복수의 속성값을 취하며 이러한 속성값에 따라 매니지먼트 에이전트가 조정과 의사결정을 하도록 한다.

Key Words : Naive Bayesian Network, Multi-Agent

1. 서 론

멀티 에이전트 시스템이란 여러 에이전트가 자율적이고 협조적인 행동에 의해 문제해결을 실현하는 시스템이다. 이러한 에이전트들이 상호 연동하여 공통의 목적을 수행하려고 할 경우에는 에이전트들을 관리하는 매니지먼트 에이전트가 요구된다[1,2,3]. 또한 에이전트를 구성하는 지식의 경우 주어진 환경을 완전하게 표현하기에는 한계가 있다. 따라서 주어진 환경에서 획득한 제한된 지식을 효율적으로 이용할 수 있는 방법이 필요하다.

본 논문에서 제안한 NBMA는 네이브 베이즈 이론과 멀티 에이전트를 결합시킴으로써 에이전트 구조를 단순화 시키고 서로 다른 에이전트간의 협동을 효율적으로 구현할 수 있게 한다. 그 구성은 고유한 속성을 지닌 여러 개의 하위 에이전트와 그들을 관리하는 하나의

매니지먼트 에이전트로 되어 있다. 하위 에이전트는 고유한 속성에 대한 메타지식을 가지고 자기가 인지한 것(percept)을 매니지먼트 에이전트에게 알려준다. 매니지먼트 에이전트는 하위 에이전트의 메타지식과 목표 함수를 바탕으로 어떤 행동이 상대적으로 더 바람직한지를 결정하기 위한 추론을 한다. 하위 에이전트 각각의 속성값은 상호 조건부 독립 가정하에 복수의 속성값을 취한다. 이러한 과정에 Naive Bayes 이론을 적용하여 매니지먼트 에이전트에 속할 확률값을 구함으로써 매니지먼트 에이전트는 행동을 결정하게 된다.

2. 관련 연구

2.1 네이브 베이즈 분류기

기계학습(Machine Learning)에서 이용되는

가장 일반적인 베이지안 네트워크 모델은 네이브 베이지안 모델이다. 이 모델은 클래스 변수 V 가 루트노드(root node), 속성변수 A 가 리프노드(leaf node)이다[4].

네이브 베이즈 분류기는 속성값들의 결합으로 표현되는 각각의 인스턴스(instance) x 와 목표 함수 $F(x)$ 로부터 학습하여 인스턴스가 해당 클래스로 분류하는데 사용된다.

간단한 베이지안 분류기(Simple Bayesian Classifier)라고도 불려지는 네이브 베이즈 분류기는 속성들이 독립이라는 가정을 하고 있다. 이러한 가정은 명백히 현실 세계에 위배되는 사실임에도 불구하고 속성들이 의존관계에 있는 많은 영역(domain)에서도 뛰어난 분류 성능을 가지고 있다[5].

새로운 인스턴스를 분류하기 위한 베이지안 접근법은 가장 큰 확률값인 v_{MAP} 을 할당하는 것이다. 속성값 $\langle a_1, a_2, \dots, a_n \rangle$ 이 주어졌을 때 다음과 같이 인스턴스를 표현한다.

$$v_{MAP} = \arg \max_{v_j \in V} P(v_j | a_1, a_2, \dots, a_n) \quad (1)$$

베이지안 이론을 사용하여 다음과 같이 표현할 수 있다.[6,7]

$$\begin{aligned} v_{MAP} &= \arg \max_{v_j \in V} \frac{P(a_1, a_2, \dots, a_n | v_j) P(v_j)}{P(a_1, a_2, \dots, a_n)} \\ &= \arg \max_{v_j \in V} P(a_1, a_2, \dots, a_n | v_j) P(v_j) \end{aligned} \quad (2)$$

샘플 데이터로부터 각 목표값의 빈도수에 의해 각각의 $P(v_j)$ 의 값을 쉽게 구할 수 있다.

네이브 베이즈 분류기는 속성값간에 상호 조건부 독립이라는 가정을 하고 있다. 이 가정에 의해 목표값이 주어졌을 때 각각 개별적인 속성들의 확률값을 곱한 것이다.

$$P(a_1, a_2, \dots, a_n | v_j) = \prod_{i=1}^n P(a_i | v_j) \quad (3)$$

식 (3)을 이용하여 식 (2)를 다음과 같이 나타낼 수 있다.

$$v_{NB} = \arg \max_{v_j \in V} P(v_j) \prod_{i=1}^n P(a_i | v_j) \quad (4)$$

따라서, 네이브 베이즈 학습 방법은 단순히 샘플 데이터의 빈도수(frequency)를 바탕으로 $P(v_j)$ 항과 $P(a_i | v_j)$ 항을 구하는 단계를 거친다.

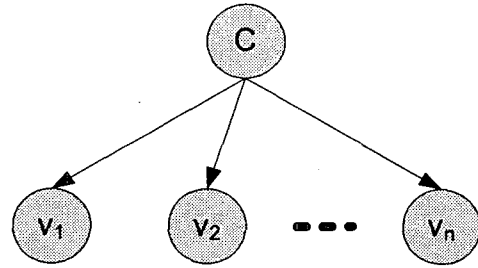


그림 1. 네이브 베이지안 네트워크 모델

3. 네이브 베이즈를 이용한 매니지먼트 에이전트(NBMA)

NBMA는 멀티 에이전트 환경을 기반으로 하는 매니지먼트 에이전트이다. 시스템의 구성은 하나의 매니지먼트 에이전트와 여러 개의 하위 에이전트들로 구성된다.

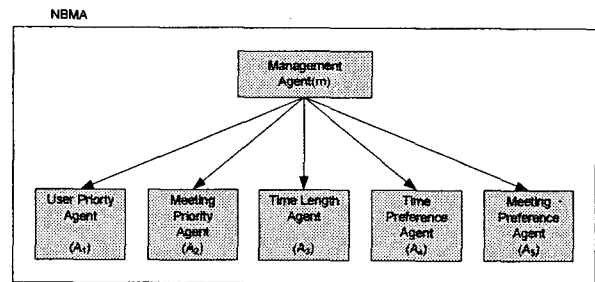


그림 2. NBMA의 구조

매니지먼트 에이전트는 하위 에이전트 $A = \{A_1, A_2, A_3, A_4, A_5\}$ 를 가진다고 한다. 각각의 하위 에이전트 A_j 는 그들의 지식을 $A_j = a_{jk} (k=1, \dots, n)$ 와 같이 복수의 속성값을 취할 수 있는 것으로 한다. 또 매니지먼트 에이전트는 클래스 $C = \{C_1, C_2, \dots, C_n\}$ 로 분류한다.

하위 에이전트 A_j 가 상호 조건부 독립이라는 가정하에 $C = C_i$ 인 매니지먼트 에이전트가 하위 에이전트의 속성 A 를 가질 확률은

$P(A | C_i) = \prod_j P(A_j = a_{jk} | C_i)$ 로 계산할 수 있다. 따라서 베이지안 정리를 이용하여 다음 식을 얻는다.

$$P(C_i | A) = \frac{P(C_i)}{P(A)} \prod_j P(A_j = a_{jk} | C_i) \quad (5)$$

식 (5)는 하위 에이전트가 속성 A 를 가질 때 매니지먼트 에이전트가 C_i 에 속할 확률을 부여한다. 즉, 하위 에이전트들의 속성 A 를 기초로 하여 매니지먼트 에이전트는 특정 클래스로 분류될 수 있다.

여기에서 $F(A)$ 는 상수로서 판별 함수 (discriminant function)를 다음과 같이 나타낼 수 있다.

$$f(m, C_j) = F(C_j) \prod_j P(A_j = a_{jk} | C_j) \quad (6)$$

식 (6)에서 매니지먼트 에이전트가 속하는 클래스 $\mathcal{C}(m)$ 은 $\mathcal{C}(m) = \arg \max_j f(m, C_j)$ 에 의해 추정할 수 있다.

4. 모의 실험

NBMA를 이용한 미팅 참가 결정 에이전트의 경우 미팅 참가자의 중요도 에이전트(UPA : User Priority Agent), 미팅의 중요도 에이전트(MPrA : Meeting Priority Agent), 미팅 시간 에이전트(TLA : Time Length Agent), 선호하는 미팅 시간대 에이전트(TPA : Time Preference Agent), 개인 스케줄을 고려한 미팅의 선호도 에이전트(MPA : Meeting Preference Agent), 미팅에 참가 가능성 에이전트(AMPA : Attending Meeting Possibility Agent)라고 가정하고 구조는 그림 2 와 같다 [8].

각 에이전트를 UP, MPr, TL, TP, MP, AMP라는 변수로 정의한다. UP, MPr, TP, MP, AMP의 속성값으로 각각 Low, Medium, High를 가지고 TL의 속성값으로 Short, Long을 가진다고 가정한다.

실험에 사용한 데이터는 표 1.과 같다. 총 108개의 데이터 중에서 랜덤으로 선택한 30개 데이터는 테스트 데이터로 사용하였고 78개는 샘플 데이터로 사용하였다. 테스트 데이터는 음영으로 표시하였다.

30개의 테스트 데이터 중 속성값이 $\langle UP=M, MPr=L, TL=S, TP=L, MP=H \rangle$ 인 인스턴스를 입력값으로 주어지면 식 (6)으로부터 각 클래스로 분류되는 확률값을 계산할 수 있다.

먼저, 주어진 속성값에 의해 AMP가 L로 분류되는 확률값은 다음과 같다.

$$\frac{P(AMP=L) P(UP=M|AMP=L) P(MPr=L|AMP=L) P(TL=S|AMP=L) P(TP=L|AMP=L) P(MP=H|AMP=L)}{\approx 0.00077468}$$

AMP가 M으로 분류되는 확률값은 다음과 같다.

$$\frac{P(AMP=M) P(UP=M|AMP=M) P(MPr=L|AMP=M) P(TL=S|AMP=M) P(TP=L|AMP=M) P(MP=H|AMP=M)}{\approx 0.0038981}$$

AMP가 H로 분류되는 확률값은 다음과 같다.

$$\frac{P(AMP=H) P(UP=M|AMP=H) P(MPr=L|AMP=H) P(TL=S|AMP=H) P(TP=L|AMP=H) P(MP=H|AMP=H)}{\approx 0.0025013}$$

표 1. 미팅 참가 결정 에이전트의 데이터

| | UP | MPr | TL | TP | MP | AMP | | UP | MPr | TL | TP | MP | AMP |
|----|----|-----|----|----|----|-----|-----|----|-----|----|----|----|-----|
| 1 | L | L | Lq | L | L | L | 55 | M | M | S | L | L | M |
| 2 | L | L | Lq | L | H | L | 56 | M | M | S | L | H | H |
| 3 | L | L | Lq | M | L | L | 57 | M | M | S | M | L | M |
| 4 | L | L | Lq | M | L | L | 58 | M | M | S | M | L | M |
| 5 | L | L | Lq | H | L | L | 59 | M | M | S | M | L | M |
| 6 | L | L | Lq | H | H | M | 60 | M | M | S | H | H | H |
| 7 | L | L | S | L | L | L | 61 | M | H | Lq | L | L | H |
| 8 | L | L | S | L | H | M | 62 | M | H | Lq | L | H | H |
| 9 | L | L | S | M | L | L | 63 | M | H | Lq | M | L | H |
| 10 | L | L | S | M | H | M | 64 | M | H | Lq | M | H | H |
| 11 | L | L | S | H | L | M | 65 | M | H | Lq | H | L | H |
| 12 | L | L | S | H | L | M | 66 | M | H | Lq | H | L | H |
| 13 | L | M | Lq | L | L | L | 67 | M | H | S | L | L | H |
| 14 | L | M | Lq | L | H | L | 68 | M | H | S | L | H | H |
| 15 | L | M | Lq | M | L | L | 69 | M | H | S | M | L | H |
| 16 | L | M | Lq | M | L | L | 70 | M | H | S | M | H | H |
| 17 | L | M | Lq | H | L | L | 71 | M | H | S | M | L | H |
| 18 | L | M | Lq | H | H | M | 72 | M | H | S | H | H | H |
| 19 | L | M | S | L | L | L | 73 | H | L | Lq | L | L | L |
| 20 | L | M | S | L | H | M | 74 | H | L | Lq | M | L | L |
| 21 | L | M | S | M | L | M | 75 | H | L | Lq | M | L | L |
| 22 | L | M | S | M | H | M | 76 | H | L | Lq | M | H | M |
| 23 | L | M | S | H | L | M | 77 | H | L | Lq | H | L | M |
| 24 | L | M | S | H | H | H | 78 | H | L | Lq | H | H | H |
| 25 | L | H | Lq | L | L | M | 79 | H | L | S | L | L | M |
| 26 | L | H | Lq | L | H | M | 80 | H | L | S | L | H | M |
| 27 | L | H | Lq | M | L | M | 81 | H | L | S | L | L | M |
| 28 | L | H | Lq | M | H | M | 82 | H | L | S | M | H | H |
| 29 | L | H | Lq | M | H | M | 83 | H | L | S | M | H | H |
| 30 | L | H | Lq | H | H | M | 84 | H | L | S | H | H | H |
| 31 | L | H | S | L | L | H | 85 | H | M | Lq | L | L | M |
| 32 | L | H | S | L | H | M | 86 | H | M | Lq | L | H | M |
| 33 | L | H | S | L | L | M | 87 | H | M | Lq | M | L | M |
| 34 | L | H | S | M | H | H | 88 | H | M | Lq | M | H | H |
| 35 | L | H | S | H | L | M | 89 | H | M | Lq | M | H | H |
| 36 | L | H | S | H | L | M | 90 | H | M | Lq | H | H | H |
| 37 | M | L | Lq | L | L | L | 91 | H | M | S | L | L | M |
| 38 | M | L | Lq | L | H | L | 92 | H | M | S | L | H | H |
| 39 | M | L | Lq | M | L | L | 93 | H | M | S | M | H | H |
| 40 | M | L | Lq | M | H | M | 94 | H | M | S | M | H | H |
| 41 | M | L | Lq | M | L | L | 95 | H | M | S | H | L | H |
| 42 | M | L | Lq | H | H | M | 96 | H | M | S | H | H | H |
| 43 | M | L | S | L | L | L | 97 | H | H | Lq | L | L | H |
| 44 | M | L | S | L | L | L | 98 | H | H | Lq | L | L | H |
| 45 | M | L | S | M | L | M | 99 | H | H | Lq | M | L | M |
| 46 | M | L | S | M | H | M | 100 | H | H | Lq | M | H | H |
| 47 | M | L | S | H | L | M | 101 | H | H | Lq | H | L | H |
| 48 | M | L | S | H | H | H | 102 | H | H | Lq | H | H | H |
| 49 | M | M | Lq | L | L | L | 103 | H | H | S | L | L | H |
| 50 | M | M | Lq | L | H | M | 104 | H | H | S | L | H | H |
| 51 | M | M | Lq | M | L | L | 105 | H | H | S | M | L | H |
| 52 | M | M | Lq | M | H | M | 106 | H | H | S | M | H | H |
| 53 | M | M | Lq | H | L | M | 107 | H | H | S | M | H | H |
| 54 | M | M | Lq | H | H | M | 108 | H | H | S | H | H | H |

실험 결과로 얻은 가장 큰 확률값은 AMP가 M으로 분류되는 확률값이며 속성값이 $\langle UP=M, MPr=L, TL=S, TP=L, MP=H \rangle$ 이 주어졌을 때 AMP가 M으로 분류된다는 것을 의미한다. 즉, 미팅 참가자의 중요도가 Medium이고 미팅의 중요도가 Low, 미팅 시간이 Short, 선호하는 미팅 시간대가 Low, 개인 스케줄을 고려한 미팅의 선호도가 High라면 미팅에 참가할 가능성이 Medium으로 분류된다는 것을 알 수 있다.

표 2.는 30개의 테스트 데이터에 대한 확률값이다. 고유한 속성을 가진 각각의 인스턴스가 각 클래스로 분류될 확률값을 보여주고 있다. 음영으로 표시한 가장 큰 확률값을 가지는 클래스를 매니지먼트 에이전트가 취할 행동으로 본다.

표 2. 테스트 데이터의 확률값

| | Low | Medium | High |
|-----|------------|-----------|------------|
| 1 | 0.0026194 | 0.0050676 | 0.00011245 |
| 4 | 0.0045319 | 0.0058066 | 0.00030235 |
| 8 | 0.0013944 | 0.0050676 | 0.00057722 |
| 12 | 0.0003486 | 0.0061937 | 0.00057722 |
| 16 | 0.0020599 | 0.0053599 | 0.00090706 |
| 19 | 0.0027466 | 0.0049896 | 0.00070841 |
| 22 | 0.00047537 | 0.0057172 | 0.0019048 |
| 25 | 0 | 0.0023389 | 0.00071216 |
| 29 | 0 | 0.0028586 | 0.00071216 |
| 33 | 0 | 0.0030492 | 0.0016451 |
| 36 | 0 | 0.0028586 | 0.0036557 |
| 41 | 0.0036367 | 0.0047644 | 0.00048727 |
| 44 | 0.00077468 | 0.0038981 | 0.0025013 |
| 49 | 0.0066121 | 0.0035983 | 0.0014618 |
| 54 | 0.00038147 | 0.004123 | 0.0035733 |
| 59 | 0.00038147 | 0.0046911 | 0.0030698 |
| 62 | 0 | 0.0016867 | 0.0075436 |
| 66 | 0 | 0.0020615 | 0.0075436 |
| 71 | 0 | 0.0023455 | 0.0064806 |
| 74 | 0.0013428 | 0.0029236 | 0.0013743 |
| 78 | 0.00033569 | 0.0035733 | 0.0013743 |
| 81 | 0.0010071 | 0.0040656 | 0.0012988 |
| 86 | 0.00061035 | 0.0026987 | 0.004123 |
| 89 | 0.00066121 | 0.0035183 | 0.0016867 |
| 93 | 0.00045776 | 0.0037529 | 0.0038983 |
| 95 | 0.00015259 | 0.0037529 | 0.003542 |
| 99 | 0 | 0.0017592 | 0.0039169 |
| 102 | 0 | 0.0016492 | 0.0087041 |
| 105 | 0 | 0.0018764 | 0.0082254 |
| 107 | 0 | 0.0018764 | 0.0074777 |

5. 결 론

에이전트들이 상호 연동하여 공통의 목적을 수행하려고 할 경우에는 에이전트들을 관리하는 매니지먼트 에이전트가 요구되며 또한 에이전트를 구성하는 지식의 경우 주어진 환경을 완전하게 표현하기에는 한계가 있다. 따라서 주어진 환경에서 획득한 제한된 지식을 효율적으로 이용할 수 있는 방법이 필요하다.

본 논문에서 제안한 NBMA는 고유한 속성을 지닌 여러 개의 하위 에이전트와 그들을 관리하는 하나의 매니지먼트 에이전트로 구성되어 있다. 하위 에이전트는 고유한 속성에 대한 메타지식을 가지고 있으며 매니지먼트 에이전트는 하위 에이전트의 메타지식을 이용한다. 하위 에이전트 각각의 속성값은 상호 조건부 독립 가정하에 복수의 속성값을 취하며 이 과정에 Naive Bayes 이론을 적용하여 매니지먼트 에이전트에 속할 확률값을 구함으로써 매니지먼트 에이전트는 행동을 결정하게 된다.

6. 참고문헌

[1] Stuart J. Russell, Peter Norvig, "Artificial

Intelligence A Modern Approach", Prentice Hall, 2003.

[2] Tomas D. Ndotsse, "Distributed Fuzzy Agent: A Framework for Intelligent Network Monitoring", 1997 IEEE International Conference on Vol. 2, pp.867 - 871, 1997.

[3] Edmund H. Durfee, Jeffrey S. Rosenschein, "Distributed Problem Solving and Multi-Agent Systems: Comparisons and Example", Proc. of Thirteen International Distributed AI Workshop, pp.94-104, 1994.

[4] Tom M. Mitchell, "Machine Learning", McGraw-Hill, 1997.

[5] Pedro Domingos, Michael Pazzani, "Beyond Independence: Conditions for The Optimality of the Simple Bayesian Classifier", International Conference on Machine Learning, 1996.

[6] David Heckerman, A Tutorial on Learning Bayesian Networks, Microsoft Research Advanced Technology Division Microsoft Corporation One Microsoft Way Redmond, pp.1-40, 1995.

[7] Nils J. Nilsson, Artificial Intelligence : A New Synthesis, Morgan Kaufmann Publisher, 1998.

[8] Chang-Shing Lee, Chen-Yu Pan, "An Intelligent Fuzzy Agent for Meeting Scheduling Decision Support System", Fuzzy Sets and Systems, Vol. 142(3) pp467-488, 2004.