



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0053180

(43) 공개일자 2015년05월15일

(51) 국제특허분류(Int. Cl.)

H04L 12/28 (2006.01)

(21) 출원번호 10-2013-0135040

(22) 출원일자 2013년11월07일

심사청구일자 없음

(71) 출원인

경희대학교 산학협력단

경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)

(72) 발명자

홍충선

경기 용인시 수지구 상현로 30-10, 233동 101호 (상현동, 성원3차상떼빌아파트)

엠디 고람 라비 올 알람

경기도 용인시 기흥구 서그내로 53번길 12-1, 102호

(74) 대리인

서재승

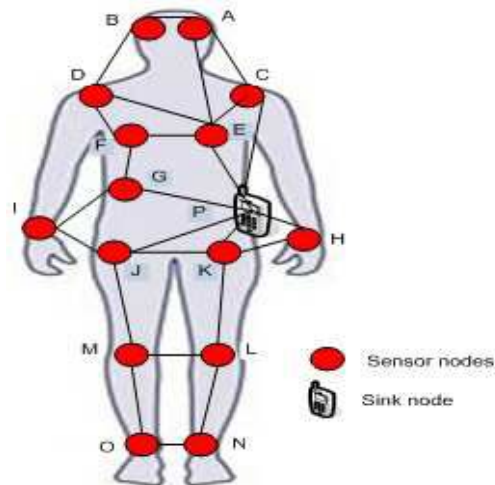
전체 청구항 수 : 총 1 항

(54) 발명의 명칭 강화학습 기반의 에너지 효율과 평형을 고려한 무선 인체 영역 네트워크 라우팅 프로토콜

(57) 요약

본 발명은 무선 신체 영역 네트워크의 고유 특성에 맞는 강화 학습을 기반으로 하는 라우팅 프로토콜로, 더욱 효율적인 전력 소비와 더불어 전력 소비의 적절한 배분이 가능한 라우팅 방법에 관한 것이다.

대표도 - 도1



명세서

청구범위

청구항 1

복잡한 연산량을 요구하지 않으며 최적에 가까운 정책을 생성하는, 센서 노드로부터 싱크 노드로의 라우팅 경로를 검색할 수 있는 강화 학습 기법의 Q-학습 방법

발명의 설명

기술 분야

[0001] 본 발명은 무선 신체 영역 네트워크의 고유 특성에 맞는 강화 학습을 기반으로 하는 라우팅 프로토콜로, 더욱 효율적인 전력 소비와 더불어 전력 소비의 적절한 배분이 가능한 라우팅 방법에 관한 것이다.

배경 기술

[0002] 무선 신체 영역 네트워크(WBAN)이 심한 감쇠, 작은 크기의 생체 센서, 높은 신뢰성과 에너지 효율 요구 및 높은 처리량, 낮은 잠복(latency) 요구 등과 같은 독특한 특징을 가진다. 이러한 요구로 인하여, 무선 신체 영역 네트워크를 위한 네트워크 프로토콜에 대한 설계가 필요하다.

발명의 내용

해결하려는 과제

[0003] 본 발명의 목적은 강화 학습(reinforcement learning) 기법의 Q-학습 방법을 사용하여 무선 신체 영역 네트워크에서 낮은 패킷 삭제 비율을 유지하면서 에너지 효율, 에너지 균형 및 최소 홉라우팅을 달성하는 것이다.

과제의 해결 수단

[0004] 복잡한 연산량을 요구하지 않으며 최적에 가까운 정책을 생성하는, 센서 노드로부터 싱크 노드로의 라우팅 경로를 검색할 수 있는 강화 학습 기법의 Q-학습 기법인 것을 특징으로 한다.

발명의 효과

[0005] 본 발명은 강화 학습(reinforcement learning) 기법의 Q-학습 방법을 사용하여 무선 신체 영역 네트워크에서 낮은 패킷 삭제 비율을 유지하면서 에너지 효율, 에너지 균형 및 최소 홉 라우팅을 달성할 수 있다.

도면의 간단한 설명

[0006] 도 1은 본 발명에 따른 무선 신체 영역 네트워크를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0007] 무선 센서 네트워크의 기존 라우팅 방법은 자원이 제한된 무선 신체 네트워크에 적합하지 않다. 따라서 복잡한 계산 없이 최적에 가까운 정책을 생성할 수 있는 몇 가지 추정 방법이 필요하다. 확률 모델을 모른 채 최적의 정책을 학습하기 때문에 강화 학습 전략을 사용하여 상태 동작 쌍으로 작업 환경을 학습하는 Q-학습 방식을 사용했다. Q-테이블과 같은 학습을 통해 최적의 정책을 사용하여 환경에 최적의 행동을 수행할 수 있다.

[0008] 신체 영역 네트워크에 데이터 패킷의 라우팅을 위한 Q-학습을 적용하려면, 센서 노드들 각각의 상태를 고려하여야 한다. 따라서 상태의 집합은 $\{X1, X2, \dots, XN\} = \{A, B, \dots, P\}$ 이다. 각 상태에서 노드는 패킷 포워딩 작업을 수행할 수 있으며 패킷을 포워딩하기 위한 또 다른 경로를 검색할 수 있다. 데이터 패킷은 성공적으로 포워딩되거나 또는 노드가 동일한 경로 또는 다른 적절한 경로를 통해 패킷을 재전송하는 경우 패킷 실패가 발생할 수 있다. 그래서, 작업 세트 $\{A1, A2, \dots\}$ 를 고려한다.

[0009] Q-학습은 상태-동작 쌍 $Q(x, a)$ 로 동작한다. 정책(π)을 사용하는 경우 상태(x)에서 취해진 동작의 값을 정책

(π)에서 취한 동작에 대한 기대 보상으로 아래의 수학적식(1)로 규정할 수 있다.

[수학적식 1]

$$Q^{\pi}(x, a) = E_{\pi}\{R_t | x_t = x, a_t = a\}$$

여기서 R_t 는 시간 t 에서의 기대 보상을 의미한다. 보상은 이전에 성공적으로 수행된 단계를 다시 수행하는 동기를 부여하는 것의 하나이다. 일반적으로, 학습 에이전트에 대한 보상으로 성공적인 작동 또는 효율적인 운영을 위해 양수 값을 할당한다. 때로는 보상이 다수의 효율적인 동작에 종속하는 다중 함수 결과를 이용하여 보상이 결정된다. 비효율적인 작업에 대해, 부정적인 보상을 제공함으로써 학습 에이전트를 처벌할 수 있다. 따라서 보상 함수의 목적은 만약 수학적식(4)에서 규정하는 신호 세기와 휴리스틱(heuristic) 거리를 고려하여 이전 시간에 패킷을 성공적으로 송신하였다면, Q-학습 에이전트로 하여금 동일한 링크를 다시 따르도록 동기를 부여하여 것이다. 또한 수학적식(8)의 전송 실패에 대해서는 에이전트에게 처벌을 적용할 수 있다. 수학적식(2)와 같이 최적의 정책에 따라 반복에 의해 근사화되는 최적 값을 검색할 수 있다.

[수학적식 2]

$$Q(x_t, a_t) \leftarrow (1 - \beta)Q(x_t, a_t) + \beta [R_t + \gamma \max_a Q(x_{t+1}, a)]$$

여기서 β 는 0 또는 1의 값을 가지는 수렴 비율이다.

신체 영역 네트워크의 효율적인 라우팅 정책을 위하여, 상태, 동작과 보상을 아래와 같이 정의한다..

상태: $X=\{x_i, \text{신체 영역 네트워크의 센서 노드}, i=(1,2,3 \dots\dots N)\}$

동작: $A=\{a(x_i|t)\}$, 센서 노드(x_i)와 센서 노드(x_j) 사이의 패킷 포워딩 동작, 여기서 x_j 는 x_i 의 송신 범위에 위치한다}

보상 함수: 보상 함수는 아래의 수학적식(3)과 같다.

[수학적식 3]

$$R_t = P_a^s(x_i, x_j)R_a^s(x_i, x_j) + (1 - P_a^s(x_i, x_j))R_a^u(x_i, x_j)$$

여기서 $P_a^s(x_i, x_j)$ 는 노드 x_i 에서 노드 x_j 로 동작 a 를 이용하여 성공적으로 패킷을 송신할 확률을 의미하며, x_i 는 x_j 에 인접하여 위치한다.

$P_a^s(x_i, x_j)$ 는 아래의 수학적식(4)와 같이 정의된다.

[수학적식 4]

$$R_a^s(x_i, x_j) = -\alpha_1 [C_i(energy_{residual}) + C_j(energy_{residual})] + \alpha_2 D_i(signal_{strength}) + \alpha_3 E_{jp}(heuristic_{distance})$$

여기서 A_1, A_2 및 A_3 는 잔여 에너지의 가중치 요인, 신호 세기 및 휴리스틱 거리를 의미하며 $A_1+A_2+A_3=1$ 이다.

$C(energy_{residual})$ 는 센서 노드 i 의 초기 에너지(E_{init})와 현재 잔여 에너지에 의존하는 잔여 에너지의 비용 함수이다. 성공적으로 패킷을 송신하는 경우, 송신기 노드(x_i)와 발신기 노드(x_j)는 모두 에너지를 손실한다. 노드가 적은 잔여 에너지를 가지고 있다면 패킷을 전달하는 노드가 사용하는 잔여 에너지 비용은 수학적식(5)와 같이 더 커진다. 노드는 수학적식(4)의 기대 함수에 따라 처벌될 것이다.

[0028] [수학식 5]

$$C(\text{energy}_{\text{residual}}) = \frac{1}{E_{\text{init}}} (E_{\text{init}} - E_{\text{res}})$$

[0029]

[0030] $D(\text{signal}_{\text{strength}})$ 는 송신기 노드(x_i)가 주변 노드(x_j)로부터 수신하는 신호 세기의 기대 함수를 의미하며 아래의 수학식(6)과 같다.

[0031] [수학식 6]

$$D(\text{Signal}_{\text{strength}}) = \lambda * \text{Kurtosis}(\text{RSS}(x_i, x_j)) = \frac{\sum (s_j - \bar{s})^4 / m}{(\sum (s_j - \bar{s})^2 / m)^2}$$

[0032]

[0033] $E(\text{Heuristic}_{\text{distance}})$ 는 싱크 노드(x_p)와 기대 수신기 노드(x_j) 사이의 직선 라인 거리의 허용가능한 휴리스틱 값에 대한 기대 함수이며, 아래의 수학식(7)과 같다.

[0034] [수학식 7]

$$E(\text{Heuristic}_{\text{distance}}) = 1 - \left[\min_{\text{dist}} h_{\text{dist}}(x_p, x_j) / \max_{\text{dist}} h_{\text{dist}}(x_p, x_j) \right]$$

[0035]

[0036] $R_a^u(x_i, x_j)$ 는 실패한 패킷 전송에 대한 보상 함수이며, 아래의 수학식(8)과 같다.

[0037] [수학식 8]

$$R_a^u(x_i, x_j) = -\kappa - \alpha_1 [C_i(\text{energy}_{\text{residual}})] + \alpha_2 D_j(\text{signal}_{\text{strength}}) + \alpha_3 E_{j,p}(\text{heuristic}_{\text{distance}})$$

[0038]

[0039] 무선 신체 영역 네트워크(WBAN)는 통신 기술, 생체 의공학, 유비쿼터스 컴퓨팅 및 에너지 효율, 신뢰성, 유비쿼터스 센서 소형화, 헬스케어 애플리케이션을 위한 나노 통신이 결합되는 새로운 무선 네트워킹 기술이다. 강화 학습에 기반한 본 발명의 라우팅은 무선 신체 센서 네트워크 또는 신체 영역 네트워크 또는 바이오 센서 네트워크와 같은 무선 개인 영역 네트워크에 적절하게 사용될 수 있는데, 네트워크 수명을 증가시키기 위해 에너지 균형과 에너지 효율을 고려하여야 한다. 따라서 본 발명은 환자 모니터링, e-헬스케어, U-헬스케어, m-헬스케어, i-헬스케어와 같은 헬스케어 애플리케이션에 적용 가능하며 운동 트레이닝 및 개인 생활 보호에도 적용 가능하다.

[0040] 도 1은 본 발명에 따른 무선 신체 영역 네트워크를 설명하기 위한 도면이다.

[0041] 도 1을 참고로 보다 구체적으로 살펴보면, 무선 신체 영역 네트워크에서 센서 노드 및 싱크 노드의 2가지 타입의 센서 노드를 고려할 수 있다. 노드는 자신의 통신 범위에 따라 서로 연결되어 있다. 도 1에 n개의 센서 노드들이 존재한다. 센서 노드의 세트로 $X = \{A, B, C, D, \dots, P\}$ 이다. 여기서 P는 싱크 노드이며 다른것들은 다른 센서 노드들이다. WBAN에서 싱크 노드로 모든 센서 노드들은 데이터 패킷을 포워드하며 싱크 노드는 무선 개인 영역 네트워크(WPAN) 또는 무선 도시 영역 네트워크(WMAN)로 데이터 패킷을 송신한다. 이웃 노드의 세트는 수신 신호의 세기에 기초하여 다른 노드들과 통신할 수 있는 노드로 규정된다. 노드는 한계 범위 내에서 통신 요구에 따라 소인 에너지를 능동적으로 조절할 수 있다.

[0042] 여기에서 복잡한 연산량을 요구하지 않으며 최적에 가까운 정책을 생성하는, 센서 노드로부터 싱크 노드로의 라우팅 경로를 검색할 수 있는 강화 학습 기법의 Q-학습 기법을 사용한다.

[0043] 본 발명에서 각 센서 노드는 Q값을 갱신하여 싱크 노드로의 거의 최단의 라우트를 학습할 수 있다. 최적 Q값에 대한 보상 함수는 잔여 에너지, RSS 및 허용 가능한 휴리스틱 거리를 보상한다. 따라서 라우팅 프로토콜(QRP)

에 기반한 본 발명의 Q-학습은 WBAN 네트워크 센서의 균형된 에너지 소비를 이용하여 낮은 패킷 손실율을 유지하며 효율적으로 최적 라우트를 선택할 수 있다.

도면

도면1

