



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2021년08월11일  
(11) 등록번호 10-2289471  
(24) 등록일자 2021년08월06일

(51) 국제특허분류(Int. Cl.)  
HO4W 74/08 (2019.01) GO6N 20/00 (2019.01)  
HO4W 72/12 (2009.01)  
(52) CPC특허분류  
HO4W 74/0808 (2013.01)  
GO6N 20/00 (2021.08)  
(21) 출원번호 10-2020-0011583  
(22) 출원일자 2020년01월31일  
심사청구일자 2020년01월31일  
(65) 공개번호 10-2021-0097955  
(43) 공개일자 2021년08월10일  
(56) 선행기술조사문헌  
Fair and Efficient Channel observation-based Listen-Before Talk (CoLBT) for LAA-WiFi Coexistence in Unlicensed LTE (2018.7.6.)\*  
(뒷면에 계속)

(73) 특허권자  
세종대학교산학협력단  
서울특별시 광진구 능동로 209 (군자동, 세종대학교)  
(72) 발명자  
김형석  
서울특별시 광진구  
라시드 알리  
서울특별시 광진구 능동로 209 세종대학교  
정철  
서울특별시 광진구 능동로 209 세종대학교  
(74) 대리인  
김현승

전체 청구항 수 : 총 6 항

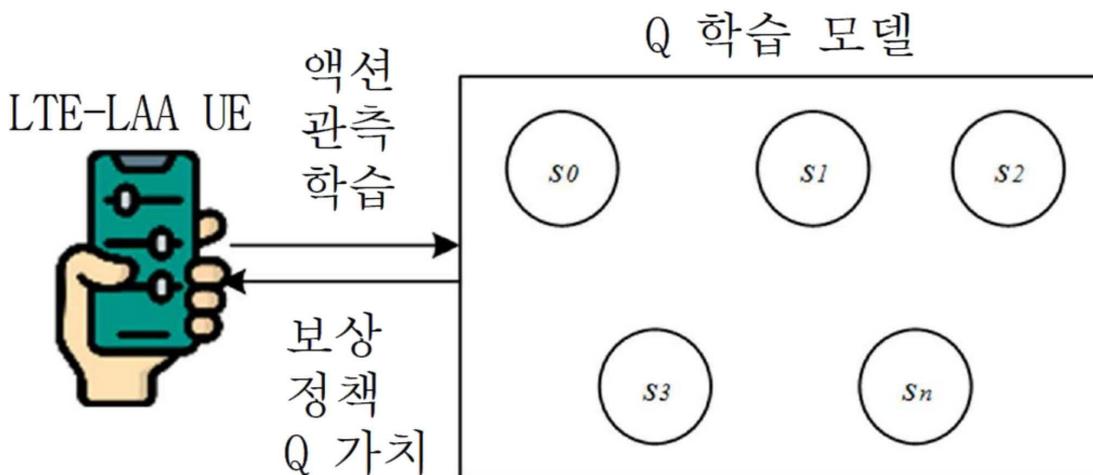
심사관 : 유환욱

(54) 발명의 명칭 강화 학습 기반의 채널 액세스 제어

(57) 요약

롱텀 에볼루션 인가 보조 액세스(Long Term Evolution Licensed Assisted Access: LTE-LAA) 및 와이파이(Wireless Fidelity: Wi-Fi)의 공존을 위한 대화 전 청취(Listen Before Talk: LBT) 메커니즘을 LTE-LAA 노드에서 구현하는 방법이 제공된다. 개시된 방법은, 확장된 클리어 채널 평가(Extended Clear Channel Assessment: ECCA)의 현재 스테이지를 위해 ECCA 경쟁 윈도우 값에 의해 정해진 범위로부터 랜덤 백오프 값을 선택하는 단계와, 비인가 대역 내의 채널의 관측에 기반하여 충돌 확률을 계산하는 단계와, ECCA의 후속 스테이지를 위해 ECCA 경쟁 윈도우 값을 조절하는 액션 정책을 학습하기 위한 강화 학습(Reinforcement Learning: RL)에서 채널의 관측에 기반한 충돌 확률을 사용하는 단계를 포함한다.

대표도



(52) CPC특허분류  
*H04W 72/1215* (2013.01)

(56) 선행기술조사문헌  
 Machine Learning Based Scheme for Contention  
 Window Size Adaptation in LTE-LAA  
 (2017.10.13.)\*  
 Q-Learning Based Fair and Efficient  
 Coexistence of LTE in Unlicensed Band  
 (2019.6.28.)\*  
 KR1020170020810 A  
 KR101845398 B1

KR1020140052704 A  
 \*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업  
 과제고유번호 1711098512  
 부처명 과학기술정보통신부  
 과제관리(전문)기관명 한국연구재단  
 연구사업명 집단연구지원(R&D)  
 연구과제명 Tactile Internet을 위한 초저지연 고신뢰 햅틱 데이터 통신 기술 및 촉각 센서/액  
 튜에이터 기술 연구  
 기여율 70/100  
 과제수행기관명 세종대학교  
 연구기간 2019.09.01 ~ 2020.02.29

이 발명을 지원한 국가연구개발사업  
 과제고유번호 1711094641  
 부처명 과학기술정보통신부  
 과제관리(전문)기관명 한국연구재단  
 연구사업명 개인기초연구(과기정통부)(R&D)  
 연구과제명 대규모 머신타입 통신/고신뢰 저지연 통신 성능 향상을 위한 평균 필드 게임 이론  
 및 비직교 다중 액세스 기반 기술 연구  
 기여율 30/100  
 과제수행기관명 세종대학교  
 연구기간 2019.06.01 ~ 2020.02.29

공지예외적용 : 있음

---

**명세서**

**청구범위**

**청구항 1**

롱텀 에볼루션 인가 보조 액세스(Long Term Evolution Licensed Assisted Access: LTE-LAA) 및 와이파이(Wireless Fidelity: Wi-Fi)에 의해 공유되는 비인가 대역(unlicensed band) 내의 채널을 다투면서 상기 LTE-LAA 및 상기 Wi-Fi의 공존을 가능하게 하는 대화 전 청취(Listen Before Talk: LBT) 메커니즘을 강화 학습(Reinforcement Learning: RL)을 기반으로 LTE-LAA 노드에서 구현하는 방법으로서, 상기 RL은 Q 학습(Q Learning: QL)을 포함하고, 상기 방법은,

확장된 클리어 채널 평가(Extended Clear Channel Assessment: ECCA)의 현재 스테이지를 위해 ECCA 경쟁 윈도우(contention window) 값에 의해 정해진 범위로부터 랜덤 백오프(random backoff) 값을 선택하는 단계 - 상기 ECCA는 상기 랜덤 백오프 값으로 백오프 카운터(backoff counter)를 시동하는 것과, 상기 현재 스테이지에서의 시간 슬롯 동안에 상기 채널이 유희(idle)인 것으로 감지되면 상기 백오프 카운터를 점감하는 것을 수반하되, 상기 LTE-LAA 노드는 상기 백오프 카운터가 만료하면 상기 채널 상에서 송신을 수행하도록 구성됨 - 와,

상기 송신에 대한 피드백 내의 부정 확인응답(Negative ACKnowledgment: NACK)의 개수를 셸프로써 의사 충돌 확률(pseudo collision probability)을 계산하는 단계와,

상기 QL에서 상기 계산된 의사 충돌 확률을 사용하여 상기 ECCA의 후속 스테이지를 위해 상기 ECCA 경쟁 윈도우 값을 조절하는 액션 정책을 학습하는 단계와,

제1 확률로 이용 국면(exploitation phase)에 진입하기로 또는 제2 확률로 탐색 국면(exploration phase)에 진입하기로 결정하는 단계를 포함하되, 상기 제1 확률 및 상기 제2 확률의 합은 1이고, 상기 이용 국면에서는 상기 후속 스테이지를 위해 상기 학습된 액션 정책에 따라 상기 ECCA 경쟁 윈도우 값이 조절되고, 상기 탐색 국면에서는 상기 후속 스테이지를 위해 상기 학습된 액션 정책과는 상관없이 상기 계산된 의사 충돌 확률에 기반하여 상기 ECCA 경쟁 윈도우 값이 확대되거나 축소되는,

방법.

**청구항 2**

제1항에 있어서,

상기 계산된 의사 충돌 확률을 사용하는 것은 상기 QL에서의 Q 가치(Q-value)의 갱신에 상기 계산된 의사 충돌 확률을 사용하는 것을 포함하되, 상기 QL의 보상(reward)은 상기 계산된 의사 충돌 확률에 의해 정의된,

방법.

**청구항 3**

삭제

**청구항 4**

삭제

**청구항 5**

제1항에 있어서,

상기 의사 충돌 확률은 상기 NACK의 상기 세어진 개수뿐만 아니라, 상기 현재 스테이지에서 상기 채널이 분주한(busy)것으로 감지된 슬롯의 개수 및 상기 랜덤 백오프 값으로부터 계산되는,

방법.

**청구항 6**

무선 통신 디바이스로서,

롱텀 에볼루션 인가 보조 액세스(Long Term Evolution Licensed Assisted Access: LTE-LAA) 및 와이파이(Wireless Fidelity: Wi-Fi)에 의해 공유되는 비인가 대역 내의 채널 상에서의 송신을 위해 구성된 송신기와, 프로세서와,

프로세서 실행가능 명령어가 저장된 메모리를 포함하되, 상기 프로세서 실행가능 명령어는 상기 프로세서에 의해 실행되는 경우 상기 프로세서로 하여금, 동작을 수행함으로써, 상기 채널을 다투면서 상기 LTE-LAA 및 상기 Wi-Fi의 공존을 가능하게 하는 대화 전 청취(Listen Before Talk: LBT) 메커니즘을 강화 학습(Reinforcement Learning: RL)을 기반으로 구현하게 하되, 상기 RL은 Q 학습(Q Learning: QL)을 포함하고, 상기 동작은,

확장된 클리어 채널 평가(Extended Clear Channel Assessment: ECCA)의 현재 스테이지를 위해 ECCA 경쟁 윈도우 값에 의해 정해진 범위로부터 랜덤 백오프 값을 선택하는 것 - 상기 ECCA는 상기 랜덤 백오프 값으로 백오프 카운터를 시동하는 것과, 상기 현재 스테이지에서의 시간 슬롯 동안에 상기 채널이 유희인 것으로 감지되면 상기 백오프 카운터를 점감하는 것을 수반하되, 상기 송신기는 상기 백오프 카운터가 만료하면 상기 송신을 수행하도록 구성됨 - 과,

상기 송신에 대한 피드백 내의 부정 확인응답(Negative ACKnowledgment: NACK)의 개수를 셈으로써 의사 충돌 확률을 계산하는 것과,

상기 QL에서 상기 계산된 의사 충돌 확률을 사용하여 상기 ECCA의 후속 스테이지를 위해 상기 ECCA 경쟁 윈도우 값을 조절하는 액션 정책을 학습하는 것과,

제1 확률로 이용 국면에 진입하기로 결정하거나 제2 확률로 탐색 국면에 진입하기로 결정하는 것을 포함하되, 상기 제1 확률 및 상기 제2 확률의 합은 1이고, 상기 이용 국면에서는 상기 후속 스테이지를 위해 상기 학습된 액션 정책에 따라 상기 ECCA 경쟁 윈도우 값이 조절되고, 상기 탐색 국면에서는 상기 후속 스테이지를 위해 상기 학습된 액션 정책과는 상관없이 상기 계산된 의사 충돌 확률에 기반하여 상기 ECCA 경쟁 윈도우 값이 확대되거나 축소되는,

무선 통신 디바이스.

**청구항 7**

제6항에 있어서,

상기 계산된 의사 충돌 확률을 사용하는 것은 상기 QL에서의 Q 가치의 갱신에 상기 계산된 의사 충돌 확률을 사용하는 것을 포함하되, 상기 QL의 보상은 상기 계산된 의사 충돌 확률에 의해 정의된,

무선 통신 디바이스.

**청구항 8**

삭제

**청구항 9**

삭제

**청구항 10**

제6항에 있어서,

상기 의사 충돌 확률은 상기 NACK의 상기 세어진 개수뿐만 아니라, 상기 현재 스테이지에서 상기 채널이 분주한 것으로 감지된 슬롯의 개수 및 상기 랜덤 백오프 값으로부터 계산되는,

무선 통신 디바이스.

**발명의 설명**

**기술 분야**

[0001] 본 개시는 강화 학습(Reinforcement Learning: RL) 기반의 채널 액세스 제어에 관한 것으로서, 더욱 구체적으로, 롱텀 에볼루션 인가 보조 액세스(Long Term Evolution Licensed Assisted Access: LTE-LAA) 및 와이파이(Wireless Fidelity: Wi-Fi)의 공존을 위해 RL을 기반으로 대화 전 청취(Listen Before Talk: LBT) 메커니즘을 구현하는 기법에 관한 것이다.

**배경 기술**

[0002] 사물 인터넷(Internet of Things: IoT)의 출현으로 이른바 연결된 디바이스(connected device)의 수가 늘어남과 인터넷을 통해 전송되는 트래픽(traffic)도 늘어났다. IoT에서는 센서(sensor)와 액츄에이터(actuator)가 대규모로 연결되므로 무선 통신이 기하급수적으로 증가된다. 5세대(5th Generation: 5G) 기술의 무선 네트워크는 수천 배의 증가된 용량(capacity)과, (가령, 수십 Gbps의 사용자당 쓰루풋(throughput) 및 1ms 미만의 지연시간(latency)를 갖는) 적어도 수억 개의 연결된 디바이스를 지원할 것으로 예상된다.

[0003] 3세대 파트너십 프로젝트(Third Generation Partnership Project: 3GPP)는 이들 연결된 디바이스를 효과적으로 지원하기 위해, 롱텀 에볼루션(Long-Term Evolution: LTE) 무선 시스템을 비인가 스펙트럼(unlicensed spectrum)에서의 활용을 위해 확장하는 것을 제안하였는데, 그 예들 중 하나가 LTE 릴리즈(Release) 13에서 도입된 인가 보조 액세스(Licensed Assisted Access: LAA) 기술이다. LTE-LAA는 5 GHz 비인가 대역을 사용한다. 그러나, 비인가 스펙트럼은 다른 무선 기술, 예를 들면 전기 전자 엔지니어 협회(Institute of Electrical and Electronic Engineers: IEEE) 802.11 무선 로컬 영역 네트워크(Wireless Local Area Network: WLAN)(이는 와이파이(Wireless Fidelity: Wi-Fi)로도 알려짐), 산업 과학 의료용(Industrial, Scientific and Medical: ISM) 공용 무선 통신 기술, 지그비(ZigBee), 블루투스(Bluetooth) 등에 의해 이미 점유되어 있다. 그래서, 이 스펙트럼은 대량의 채널 경쟁(channel contention)으로 인해 상당한 간섭에 직면할 것이다. LTE-LAA의 사용은 용량을 개선하고 원활한 사용자 경험 품질(Quality of Experience: QoE)을 달성하지만, LTE-LAA 및 앞서 언급된 상이한 무선 기술(가령, Wi-Fi)의 다양한 네트워크가 공유된 스펙트럼을 사용할 수 있게 한다는 문제가 고려될 필요가 있다. 한 가지 중요한 문제는 상이한 공존 기술 간의 간섭의 조정(coordination) 및 관리(management)이다. 특히, LTE-LAA 및 Wi-Fi의 서로 다른 물리 계층(physical layer) 및 매체 액세스 제어(Medium Access Control: MAC) 계층 구성은 이들의 효율적인 공존에서의 난관으로 이어진다. 도 1은 LTE-LAA 및 Wi-Fi가 공존하여 네트워크 환경에서 간섭이 발생한 시나리오를 예시하는 도면이다. 도 1에 도시된 예시적인 네트워크 환경(100) 내에는 셀(110)을 서빙하는(serve) 진화된 노드 B(evolved Node B: eNB)(120) 및 사용자 장비(User Equipment: UE)(123, 125, 127)와 같은 LTE-LAA 노드(node)가 존재하며, 또한 Wi-Fi 프로토콜에 따라 비인가 대역 스펙트럼을 사용하는 액세스 포인트(Access Point: AP)(130, 132) 및 스테이션(station)(STA)(133, 135)와 같은 Wi-Fi 노드가 존재한다.

[0004] 현재, LTE-LAA는 이른바 대화 전 청취(Listen-Before-Talk: LBT) 메커니즘을 사용하고, Wi-Fi는 충돌 회피식 캐리어 감지 다중 액세스(Carrier Sense Multiple Access with Collision Avoidance: CSMA/CA)를 채널 액세스 메커니즘으로서 사용한다. Wi-Fi에서의 MAC 계층 채널 액세스의 주안점은 공정한 MAC 계층 리소스 할당(MAC layer Resource Allocation: MAC-RA) 방안을 사용하여 채널 활용을 최대화하는 것이다. MAC-RA 방안은 무선 스펙트럼을 적절하게 액세스하기 위해 분산 조정 기능(Distributed Coordination Function: DCF) 기반의 CSMA/CA 메커니즘을 사용한다. 이 메커니즘에서, 이른바 이진 지수 백오프(Binary Exponential Backoff: BEB)에 따라, Wi-Fi 스테이션(station)(STA)은 데이터 프레임의 송신 전에 채널을 관측하기 위해 경쟁 윈도우(contention window)에서 대기할 시간 슬롯(time slot)의 개수를 랜덤 값(random value)으로서 선택하되, 충돌이 감지되면 경쟁 윈도우의 가능한 최대 크기 값을 지수적으로 증가시키고 송신이 성공하면 그 값을 초기치로 재설정한다. 그리고, BEB는 송신 기회(Transmission Opportunity: TXOP)의 최대 지속시간에 제한을 부과하니, 다시 말해, Wi-Fi STA은 각 데이터 프레임을 위해서 채널을 두고 경쟁할 필요가 있다. 반면에, 표준 LTE는 최소 갭(minimum gap)을 갖는 연속 데이터 송신(continuous data transmission)을 사용한다. 이러한 구성상 차이에 따른 난관을 이겨 내기 위한 기법 중 하나가 LTE-LAA에 맞춰진 LBT 메커니즘인데, 이는 비인가 대역을 액세스하기 위해 CSMA/CA와 제법 비슷한 매체 액세스 절차를 수행한다. LTE-LAA의 LBT는 데이터 프레임을 송신하려고 시도하기 전에 채널 상에서의 에너지 레벨을 판정하는 클리어 채널 평가(Clear Channel Assessment: CCA)를 사용한다. 이러한 CCA 기반 LBT 메커니즘에서, 데이터 프레임을 송신하고자 하는 LTE-LAA 노드(가령 eNB 또는 UE)는 CCA를 위해 초기 유예 기간(deferral period) 동안 채널이 유희(idle)인지를 관측하여야 한다. 이후에, 도 2에서 볼 수 있는 바와 같이, 확장된 CCA(Extended CCA: ECCA)로 알려진 또 다른 유예 기반 절차가 수행될 수 있다. ECCA에서, 백오프 카운터(backoff counter)의 랜덤 값 b가 선택될 수 있는데, 이 값은 TXOP 전에 감지될 필요가 있는 유희 관측 슬롯의 개수를 가리키며, 각 슬롯의 지속시간(duration)(이는 이른바 CCA 슬롯의 지속시간

(duration)  $\sigma$  임) 동안 채널이 유희하다고 검출되면 백오프 카운터는 1씩 감소된다. 랜덤 백오프 값  $b$ 는  $[0, CW-1]$ 의 범위로부터 선택되는데, CW는 경쟁 윈도우의 금번 상한 값(이하에서 ECCA를 위한 경쟁 윈도우 값 또는 ECCA 경쟁 윈도우 값으로도 칭해짐)이고 지수 백오프에 따라 달라진다. 초기에, CW는 최소값  $CW_{min}$ 으로 설정된다. 충돌 검출 시에 CW는 지수적으로 증가되며, 그러한 검출이 없으면 CW는 최소값  $CW_{min}$ 으로 재설정된다.

[0005] Wi-Fi에서는 확인응답(ACKnowledgment: ACK) 제어 메시지를 사용하여 충돌 감지가 수행되는 반면, LTE에서는 그러한 프레임이 없기에 현재의 TXOP의 하이브리드 자동 재송 요청(Hybrid Automatic Repeat reQuest: HARQ) 피드백(feedback)(이는 현재의 TXOP에서의 부정 확인응답(Negative ACK: NACK)의 수를 나타냄)에 기반하여 충돌이 검사된다. 3GPP에서는, 최근의 TXOP의 HARQ 피드백에서 80%나 NACK이면 충돌로 간주하여 CW를 증가시키는 것이 제안된 바 있다. 그런데, LTE-LAA는 단일의 송신 프레임에서 여러 디바이스를 스케줄링하는바, 통상적으로 80% NACK 임계치가 충족되기는 어렵다. 그러면, TXOP 내의 송신된 데이터 프레임의 80% 미만에 대해 NACK가 있는 충돌은 무시될 것이다. 더욱이, LTE의 내재적인 지연시간으로 인해, 특정 서브프레임과 연관된 HARQ 피드백은 송신되고 나서 적어도 4ms 후에 수신된다. 이런 연유로, 3GPP는 최소한의 지연으로 CW를 갱신하기 위해, TXOP의 첫 번째 서브프레임 동안에 검출된 충돌만 고려할 것(즉, TXOP 내의 나머지 서브프레임에서의 충돌은 무시됨)을 제안한다.

[0006] 이와 같이, LTE-LAA 및 Wi-Fi 공존 시나리오에서 LBT가 채택되지만, Wi-Fi의 성능은 LBT 채널 액세스 파라미터의 구성에 크게 의존한다. 따라서, LTE-LAA와 Wi-Fi 간의 공정성을 개선하기 위한 향상된 LBT 메커니즘은 여전히 필요하다.

### 발명의 내용

#### 해결하려는 과제

[0007] 롱텀 에볼루션 인가 보조 액세스(Long Term Evolution Licensed Assisted Access: LTE-LAA) 및 와이파이(Wireless Fidelity: Wi-Fi)의 공존을 위한 강화 학습 작동식 LBT(Reinforcement Learning-enabled LBT: ReLBT) 메커니즘을 구현하는 기법이 본 문서에 개시된다.

#### 과제의 해결 수단

[0008] 적어도 하나의 실시예에 따르면, 롱텀 에볼루션 인가 보조 액세스(Long Term Evolution Licensed Assisted Access: LTE-LAA) 및 와이파이(Wireless Fidelity: Wi-Fi)의 공존을 위한 대화 전 청취(Listen Before Talk: LBT) 메커니즘을 LTE-LAA 노드에서 구현하는 방법은, 확장된 클리어 채널 평가(Extended Clear Channel Assessment: ECCA)의 현재 스테이지(stage)를 위해 ECCA 경쟁 윈도우(contention window) 값에 의해 정해진 범위로부터 랜덤 백오프(random backoff) 값을 선택하는 단계(상기 ECCA는 상기 랜덤 백오프 값으로 백오프 카운터(backoff counter)를 시동하는 것과, 상기 LTE-LAA 및 상기 Wi-Fi에 의해 공유되는 비인가 대역(unlicensed band) 내의 채널의 관측을 위해 상기 현재 스테이지에서의 시간 슬롯 동안에 상기 채널이 유희(idle)인 것으로 감지되면 상기 백오프 카운터를 점감하는(decrement) 것을 수반하되, 상기 LTE-LAA 노드는 상기 백오프 카운터가 만료하면(expire) 상기 채널 상에서 송신을 수행하도록 구성됨)와, 상기 채널의 상기 관측에 기반하여 의사 충돌 확률(pseudo collision probability)을 계산하는 단계와, 상기 ECCA의 후속 스테이지를 위해 상기 ECCA 경쟁 윈도우 값을 조절하는 액션 정책(action policy)을 학습하기 위한 강화 학습(Reinforcement Learning: RL)에서 상기 계산된 의사 충돌 확률을 사용하는 단계를 포함한다.

[0009] 상기 RL은 Q 학습(Q Learning: QL)을 포함할 수 있다.

[0010] 상기 계산된 의사 충돌 확률을 사용하는 단계는 상기 QL에서의 Q 가치(Q-value)의 갱신에 상기 계산된 의사 충돌 확률을 사용하는 단계를 포함할 수 있는데, 상기 QL의 보상(reward)은 상기 계산된 의사 충돌 확률에 의해 정의될 수 있다.

[0011] 상기 방법은, 상기 후속 스테이지를 위해 상기 학습된 액션 정책에 따라 상기 ECCA 경쟁 윈도우 값을 조절하는 이용 국면(exploitation phase) 또는 상기 RL과는 별개로 상기 계산된 의사 충돌 확률에 기반하여 상기 채널의 유희 여부에 따라 상기 후속 스테이지를 위해 상기 ECCA 경쟁 윈도우 값을 확대하거나 축소하는 탐색 국면(exploration phase) 중 하나에 진입하기로 결정하는 단계를 더 포함할 수 있다.

[0012] 상기 이용 국면 및 상기 탐색 국면은 각각 제1 확률 및 제2 확률로 진입이 결정될 수 있다(상기 제1 확률과 상

기 제2 확률의 합은 1임).

- [0013] 상기 관측은 상기 현재 스테이지에서 상기 채널의 유희 여부를 감지하는 것 및 상기 현재 스테이지에서 상기 백오프 카운터가 만료된 후에 수행된 상기 송신에 대한 피드백을 수신하는 것을 포함할 수 있다.
- [0014] 상기 의사 충돌 확률은 상기 현재 스테이지에서 상기 백오프 카운터가 만료된 후 상기 송신에 대한 부정 확인응답(Negative ACKnowledgment: NACK)의 개수와 함께, 상기 현재 스테이지에서의 상기 관측 동안에 검출된 분주한(busy) 슬롯의 개수 및 상기 랜덤 백오프 값으로부터 계산될 수 있다.
- [0015] 적어도 하나의 실시예에 따르면, 무선 통신 디바이스는, 롱텀 에볼루션 인가 보조 액세스(Long Term Evolution Licensed Assisted Access: LTE-LAA) 및 와이파이(Wireless Fidelity: Wi-Fi)에 의해 공유되는 비인가 대역 내의 채널 상에서의 송신을 위해 구성된 송신기와, 프로세서와, 프로세서 실행가능 명령어가 저장된 메모리를 포함하되, 상기 프로세서 실행가능 명령어는 상기 프로세서에 의해 실행되는 경우 상기 프로세서로 하여금, 확장된 클리어 채널 평가(Extended Clear Channel Assessment: ECCA)의 현재 스테이지를 위해 ECCA 경쟁 윈도우 값에 의해 정해진 범위로부터 랜덤 백오프 값을 선택하는 것(상기 ECCA는 상기 랜덤 백오프 값으로 백오프 카운터를 시동하는 것과, 상기 채널의 관측을 위해 상기 현재 스테이지에서의 시간 슬롯 동안에 상기 채널이 유희인 것으로 감지되면 상기 백오프 카운터를 점감하는 것을 수반하되, 상기 송신기는 상기 백오프 카운터가 만료하면 상기 송신을 수행하도록 구성됨)과, 상기 채널의 상기 관측에 기반하여 의사 충돌 확률을 계산하는 것과, 상기 ECCA의 후속 스테이지를 위해 상기 ECCA 경쟁 윈도우 값을 조절하는 액션 정책을 학습하기 위한 강화 학습(Reinforcement Learning: RL)에서 상기 계산된 의사 충돌 확률을 사용하는 것을 포함하는 동작을 수행함으로써 상기 LTE-LAA 및 상기 Wi-Fi의 공존을 위한 대화 전 청취(Listen Before Talk: LBT) 메커니즘을 구현하게 한다.
- [0016] 상기 RL은 Q 학습(Q Learning: QL)을 포함할 수 있다.
- [0017] 상기 계산된 의사 충돌 확률을 사용하는 것은 상기 QL에서의 Q 가치의 갱신에 상기 계산된 의사 충돌 확률을 사용하는 것을 포함할 수 있는데, 상기 QL의 보상은 상기 계산된 의사 충돌 확률에 의해 정의될 수 있다.
- [0018] 상기 동작은 상기 후속 스테이지를 위해 상기 학습된 액션 정책에 따라 상기 ECCA 경쟁 윈도우 값을 조절하는 이용 국면 또는 상기 RL과는 별개로 상기 계산된 의사 충돌 확률에 기반하여 상기 채널의 유희 여부에 따라 상기 후속 스테이지를 위해 상기 ECCA 경쟁 윈도우 값을 확대하거나 축소하는 탐색 국면 중 하나에 진입하기로 결정하는 것을 더 포함할 수 있다.
- [0019] 상기 이용 국면 및 상기 탐색 국면은 각각 제1 확률 및 제2 확률로 진입이 결정될 수 있다(상기 제1 확률과 상기 제2 확률의 합은 1임).
- [0020] 상기 관측은 상기 현재 스테이지에서 상기 채널의 유희 여부를 감지하는 것 및 상기 현재 스테이지에서 상기 백오프 카운터가 만료된 후에 수행된 상기 송신에 대한 피드백을 수신하는 것을 포함할 수 있다.
- [0021] 상기 의사 충돌 확률은 상기 현재 스테이지에서 상기 백오프 카운터가 만료된 후 상기 송신에 대한 부정 확인응답(Negative ACKnowledgment: NACK)의 개수와 함께, 상기 현재 스테이지에서의 상기 관측 동안에 검출된 분주한 슬롯의 개수 및 상기 랜덤 백오프 값으로부터 계산될 수 있다.
- [0022] 전술된 개요는 상세한 설명에서 추가로 후술되는 몇몇 양상을 단순화된 형태로 소개하기 위해 제공된다. 이 개요는 청구된 주제(subject matter)의 중요 특징 또는 필수적 특징을 식별하도록 의도되지 않고, 청구된 주제의 범위를 정하는 데 사용되도록 의도되지도 않는다. 나아가, 청구된 주제는 본 명세서에서 논의되는 임의의 또는 모든 이점을 제공하는 구현에 한정되지 않는다.

**발명의 효과**

- [0023] 본 발명의 실시예에 따르면, 강화 학습(Reinforcement Learning: RL)을 사용하여 LBT 액세스 프로토콜과 연관된 채널 액세스 파라미터를 지능적으로 조절하는 기법이 제공된다.
- [0024] 본 발명의 실시예는 IoT 시스템에서의 LTE-LAA 및 Wi-Fi 공존을 위한 RL 작동식 LBT(RL-enabled LBT: ReLBT) 메커니즘을 구현하는바, 이는 LBT를 위한 MAC-RA 채널 액세스 파라미터를 지능적으로 최적화한다.
- [0025] 본 발명의 실시예에 따른 ReLBT 메커니즘은 RL 접근법을 통해서 환경으로부터 상호작용(interaction) 및 관측(observation)에 의해 최적의 액션(action)을 찾는다.

[0026] 본 발명의 실시예에 따른 ReLBT 메커니즘은 RL 모델을 위해 채널 관측 기반 충돌 확률을 송신 시도의 보상(reward)으로 정의하는바, HARQ 피드백의 80% NACK을 기다리는 것 대신에, 누적된 보상을 경쟁 파라미터의 확대(scaling-up) 및 축소(scaling-down)을 위해 이용한다(exploit).

[0027] 본 발명의 실시예에 따른 ReLBT 메커니즘은 RL 접근법에 의한 채널 액세스 파라미터의 최적화의 이용 및 RL 접근법과는 무관하게 채널 관측 기반 충돌 확률에 기반하여 채널 액세스 파라미터를 조절하는 탐색을 적절히 조화시킨다.

**도면의 간단한 설명**

[0028] 도 1은 LTE-LAA 및 Wi-Fi가 공존하여 네트워크 환경에서 비인가 대역의 채널 상에서 간섭이 발생하는 시나리오를 예시하는 도면이다.

도 2는 LTE-LAA의 LBT 메커니즘의 개요를 설명하기 위한 도면이다.

도 3은 본 발명의 실시예에 따라 QL 기반 모델을 사용하는 ReLBT 메커니즘을 도식적으로 보여주는 도면이다.

도 4는 본 발명의 실시예에 따라 ReLBT 메커니즘을 구현하기 위한 프로세스의 흐름도이다.

도 5는 본 발명의 실시예에 따라 ReLBT 메커니즘을 구현하기 위한 무선 통신 장치의 블록도이다.

도 6은 본 발명의 실시예에 따른 ReLBT 메커니즘의 시물레이션을 위한 배치 시나리오를 개략적으로 도시한 도면이다.

도 7 내지 도 10은 본 발명의 실시예에 따른 ReLBT 메커니즘의 성능을 설명하기 위한 도면이다.

**발명을 실시하기 위한 구체적인 내용**

[0029] 이하에서는, 첨부된 도면을 참조하여 본 발명의 실시예를 상세히 설명한다. 본 발명은 여러 가지 실시예를 가질 수 있고, 몇몇 실시예가 본 명세서에 개시된다. 그러나, 이는 본 발명에 대한 한정이나 예시로서 제공되며, 본 발명의 사상 및 범주에 속하는 모든 변형, 균등물 내지 대체물을 망라하는 것으로 이해되어야 한다. 개시된 실시예에 따른 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 다음의 상세한 설명에서 특정한 세부사항이 제공되는데, 몇몇 실시예는 이들 세부사항 중 일부 또는 전부가 없더라도 실시될 수 있다. 또한, 본 발명의 다양한 양상을 불필요하게 모호하게 하지 않도록 공지 기술의 구체적인 설명은 생략될 수 있다.

[0030] 후술되는 용어는 단지 특정 실시예를 설명하기 위해 사용된 것으로, 한정적 의미로 고려되고자 의도된 것이 아니다. 단수 형태의 표현은 명확하게 달리 사용되지 않는 한, 복수 형태의 의미를 포함한다. 또한, 이 문서에서, "포함하다" 또는 "가지다"와 같은 용어는 어떤 특징, 숫자, 단계, 동작, 구성요소, 정보 또는 이들의 조합이 존재함을 나타내려는 것이며, 하나 또는 그 이상의 다른 특징, 숫자, 단계, 동작, 구성요소, 정보 또는 이들의 조합의 존재 또는 가능성을 배제하지 않는 것으로 이해되어야 한다.

[0031] 본 발명의 실시예에 따라 LTE-LAA 및 Wi-Fi의 공존을 위한 향상된 LBT 메커니즘을 구현하는 기법은, LTE를 위한 표준적인 HARQ 기반 충돌 검출 메커니즘 대신에, 채널 관측 기반 충돌 확률(channel observation-based collision probability)  $p_{obs}$ 를 사용하는 더욱 현실적인 메커니즘을 채택한다. 그러한 개선된 LBT 메커니즘은 강화 학습(Reinforcement Learning: RL) 접근법을 활용한다는 점에서 RL 작동식 LBT(RL-enabled LBT: ReLBT) 메커니즘으로 칭해질 수 있고, 예컨대 도 1의 네트워크 환경(100) 내에 존재하는 LTE-LAA 노드(가령, LTE-LAA eNB(120), LTE-LAA UE(123, 125, 127) 등등) 내에 구현될 수 있다. 예로서, ReLBT 메커니즘은 RL의 일종인 Q 학습(Q Learning: QL)을 기반으로 할 수 있는데, 도 3은 이 메커니즘을 위한 QL 모델을 도식적으로 보여준다. 아래에서 상세히 논의되는 바와 같이, 본 발명의 실시예에 따른 ReLBT 메커니즘은 기존의 LBT 메커니즘에 대한 약간의 수정을 요구할 뿐이므로, 그것과의 호환가능성(compatibility)을 온전히 유지할 수 있다.

[0032] 이제, 도 4를 참조하여, 예시적인 ReLBT 메커니즘을 구현하기 위한 프로세스(400)를 살펴본다. 예시적인 ReLBT 메커니즘에서, CCA 동안에 통신 매체가 유희였던 후에, 송신을 위해 채널을 다투는 LTE-LAA 노드는 랜덤 백오프 값  $b$ 를 선택함(동작(420))으로써 ECCA 절차로 나아갈 수 있다. 실시예에 따르면, CCA 직후에 오는 시간은 여러 관측 시간 슬롯  $a$ 로 구분될 수 있다. 그러한 슬롯  $a$ 의 지속시간은 유희 기간 동안에는 불변의 슬롯 시간  $\sigma$ 이거나 아니면 일정치 않은 분주(busy)(네트워크 내의 다른 디바이스에 의한 충돌되는 송신 또는 성공적인 송신) 기간이다. 동작(430)에서 기간  $\sigma$ 의 관측 시간 슬롯 동안 채널이 유희인 것으로 감지되면, LTE-LAA 노드는 랜덤

백오프 값  $b$ 를 1만큼 점감한다(decrement)(동작(434)). 동작(432)에서 랜덤 백오프 값  $b$ 가 0에 도달하여야 시간 슬롯의 시작에서 TXOP가 사용된다(동작(440)). 또한, 분주한 채널이 된 것으로 동작(430)에서 감지되면, LTE-LAA 노드는 랜덤 백오프 값  $b$ 를 동결하고(freeze)(동작(436)), 채널 감지를 계속한다. 이후에 동작(430)에서 채널이 다시 유향인 것으로 검출되면, LTE-LAA 노드는 랜덤 백오프 값  $b$ 의 카운트다운(countdown)을 재개한다(동작(434)). 실시예에서, 네트워크 내의 각각의 노드는 채널 관측 기반 충돌 확률  $p_{obs}$ 를 능히 측정할 수 있다(동작(450)). 이 측정된  $p_{obs}$ 는 그것을 추정하는 데에 각각의 노드가 관측을 통해 송신 실패(즉, NACK)의 횟수를 세고(count) 그 횟수를 송신 시도의 총 횟수로 나누는 것이 요구되므로 의사 충돌 확률(pseudo collision probability)로 칭해질 수 있다. 그런데, 실시예에 따르면, ECCA 절차의 분주 및 유향 기간을 셈으로써  $p_{obs}$ 를 위한 더 현실적인 관측을 할 수가 있다. 예시적인 ReLBT 메커니즘은 최근의 TXOP에서의 NACK의 개수  $S_{nack}$ 를 획득하고(동작(442)) 관측 시간 슬롯  $a$ 에서 분주한 슬롯의 개수  $S_b$ 를 셈(동작(436))으로써 매 ECCA 백오프 스테이지(stage)에서  $p_{obs}$ 를 갱신한다(동작(450)). 관측된 시간 슬롯  $a$ (분주 및 유향)의 총 개수는  $B_{obs} = b + S_b$ 에 의해 표현될 수 있다. 그러면, 다음과 같이, LTE-LAA 노드는 해당 ECCA 백오프 스테이지의  $B_{obs}$ 로부터 채널 관측 기반 충돌 확률  $p_{obs}$ 를 갱신할 수 있다.

**수학식 1**

$$p_{obs} = \frac{(S_b + S_{nack})}{(S_{nack} + B_{obs})}$$

[0033]

여기서  $S_b = \sum_{k=0}^{B_{obs}-1} S_k$  이고, 관측 시간 슬롯  $k$ 에 대해, 채널이 유향이면  $S_k=0$ 이고, 채널이 분주하면  $S_k=1$ 이다(즉, 분주한 채널이 감지될 때마다 동작(436)에서  $S_b$ 가 점증됨(incremented)). 예시적인 ReLBT 메커니즘은 관측된 실제적인 충돌 확률에 기반하여 ECCA 경쟁 윈도우 값을 갱신하는데, 이는 더욱 적응적인 경쟁 절차를 제공하며, 따라서 공존하는 두 기술 간의 공정한 공유를 가능하게 한다. 이와 같이, 채널의 관측을 기반으로 한 타당한 충돌 확률  $p_{obs}$ 가 예시적인 ReLBT 메커니즘에서 중요한 역할을 한다.

[0034]

구체적으로, 예시적인 ReLBT 메커니즘에 따르면, LTE-LAA 노드는 충돌을 검출하고 만일 채널 관측 기반 충돌 확률  $p_{obs}$ 가 0보다 크면 현재의 ECCA 경쟁 파라미터를 스케일링한다(scale). 이 프로세스는 설정된 현재의 TXOP의 피드백으로부터 어떤 NACK도 수신되지 않더라도, 여전히 경쟁 파라미터는 관측 동안의 분주한 슬롯으로 인해 스케일링될 수 있음을 시사한다. 기존의 HARQ 피드백 기반 충돌 검출 메커니즘과 달리  $p_{obs}$ 에 기반하여 경쟁 파라미터를 동적으로 조절하는 것은, 예컨대 80% NACK 임계치가 충족되기 어려운 경우에, 상당히 유리할 것이다. 더욱이, 예시적인 ReLBT 메커니즘은 종래의 LBT에서 CW를 지속적으로 증가시키고 최소 값으로 재설정하는 것과는 달리, 경쟁 윈도우 CW의 확대와 축소를 수행한다. 실시예에서, 만일  $p_{obs} > 0$ 인 경우(즉, 분주한 슬롯 및/또는 NACK가 존재하는 경우) CW가 확대될 수 있고, 만일  $p_{obs} = 0$ 인 경우(즉, 분주한 슬롯도 NACK도 전혀 없는 경우) CW가 축소될 수 있다. 몇몇 예시적인 구현에서, ECCA 경쟁 윈도우의 확대 및 축소는 다음과 같이 동작할 수 있다.

[0035]

**수학식 2**

$$CW_{cur} = \begin{cases} \min[2 \times CW_{pre} \times \omega^{p_{obs}}, CW_{max}], & \forall p_{obs} > 0 \\ \max\left[\frac{CW_{pre} \times \omega^{p_{obs}}}{2}, CW_{min}\right], & \forall p_{obs} = 0 \end{cases}$$

[0036]

[0037] 여기서  $CW_{cur}$  는 이전의 ECCA 경쟁 윈도우 값  $CW_{pre}$  로부터 확대 또는 축소된 금번의 ECCA 경쟁 윈도우 값이고,  $CW_{max}$  및  $CW_{min}$  는 각각 그 경쟁 윈도우의 최대값 및 최소값이며,  $\omega$  는 관측된  $p_{obs}$ 에 따라 ECCA 경쟁 윈도우 값의 크기를 적응적으로 제어하는 일정한 설계 파라미터(가령,  $\omega = CW_{min}$ 으로서 표현됨)이다.

[0038] 나아가, 예시적인 ReLBT 메커니즘은, 도 3에 도시된 바와 같이, Q 학습(Q Learning: QL)을 기반으로 경쟁 파라미터를 선택하는 것을 수반하는바, 채널 충돌이 감소되게 하고 더욱 공정한 채널 액세스를 가능하게 한다. QL은 강화 학습(Reinforcement Learning: RL) 모델 중 하나로서, 학습자가 불확실한 환경과 상호작용하여 성능을 최적화하는 문제 전체를 상당히 반영한다. 예를 들어, 특정 목표 지향적인 학습자는 LTE-LAA 환경 내에서 Wi-Fi와의 공정한 공존이라는 측면에서 그것의 성능을 최대화하고자 하는 무선 노드(가령, UE)일 수 있다.

[0039] 구체적으로, 실시예에 따르면, ReLBT 메커니즘을 위한 QL 모델은 학습자(가령, LTE-LAA UE), 환경(가령, LTE-LAA 및 Wi-Fi의 공존을 위한 네트워크 환경), 정책(가령, ECCA 경쟁 윈도우 값의 확대 및 축소), 보상(가령, 채널 관측 기반 충돌 확률  $p_{obs}$ ) 및 누적된 보상인 Q 가치 함수(Q-value function)로 구성된다. 주어진 시간에서의 학습자의 거동 및 학습은 그것이 따르는 정책에 의존한다. 한편, 정책은 환경의 인지된 상태를 맵핑하기 위해 총체적인 액션을 결정하는 규칙이다. 보상은 각 단계에서의 상황에 의해 결정되는 정량적인 값인데, QL 기반 모델에서 학습자의 유일한 목표는 장기적인 실행에 걸쳐서 누적된 보상을 최대화하는 것이다. 보상은 특정 상태에서 임의의 단일 액션에 대한 즉각적인 정량적 값인 반면, Q 가치는 그 상태에서 획득된 누적된 보상을 나타낸다. 어떤 상태가 언제나 낮은 보상을 산출하면서도, 연이어 뒤따르는 다른 상태가 높은 보상을 가져오는 까닭에 높은 Q 가치를 가질 수가 있다.

[0040] 예시적인 ReLBT 메커니즘은 상태의 세트 S를 포함한다. 이 세트 내의 임의의 상태  $s_1, s_2, \dots$ , 또는  $s_n$ 는 ECCA에서의 백오프 스테이지인데, 여기서 학습자는 허용되는 액션( $p_{obs}>0$ 이면 CW를 증가시키는 것과,  $p_{obs}=0$ 이면 CW를 감소시키는 것)의 세트 A으로부터 액션을 취할 수 있다. 특정한 상태 s에서 정책  $\pi$ 에 따라 액션 a를 수행함으로써, 학습자는 Q 가치 함수로 알려진 누적 보상  $Q(s,a)$ 를 계산하고(동작(462)) 이를 이용하기 위해, 보상  $R(s,a)$ 를 계산하고(동작(460)) 이를 축적한다.

[0041] 이하에서, 논의의 편의를 위해,  $S=\{1, 2, \dots, n\}$ 이 환경의 n개의 가능한 상태의 유한 세트를 나타내도록 하고,  $A=\{0, 1\}$ 이 허용되는 액션의 유한 세트를 나타내도록(단, 0은 CW 감소를 나타내고 1은 CW 증가를 나타냄) 하자. 예시적인 ReLBT 메커니즘에 따르면, 시간 슬롯 t에서, 학습자는 현재의 상태  $s_t$ (여기서  $s_t \in S$ )를 관측하고, 액션  $a_t$ (여기서  $a_t \in A$ )를 정책  $\pi_t$ 에 기반하여 취한다. 실시예에서, ReLBT에서의 학습자의 디폴트(default) 정책은 충돌의 경우에는 그것의 상태를 증가시키는 것이고, 성공적인 송신의 경우에는 이를 감소시키는 것이다. 따라서, 다음 수학적식에 따라 액션  $a_t$ 는 환경의 상태를  $s_t$ 로부터  $s_{t+1}$ (여기서  $s_{t+1}=s' \in S$ )로 변화시킨다.

**수학적식 3**

[0042] 
$$\pi(a|s) = \begin{cases} s' = s + 1, & \forall p_{obs} > 0 \\ s' = s - 1, & \forall p_{obs} = 0 \end{cases}$$

[0043] 예시적인 ReLBT 메커니즘에 따르면, QL 모델에서,  $Q(s,a)$ 는 누적적 보상을 추정하며 다음과 같이 갱신된다(동작(462)).

**수학적식 4**

[0044] 
$$Q(s, a) \leftarrow (1 - \gamma) \times Q(s, a) + \gamma \times \Delta Q(s, a)$$

[0045] 여기서  $\gamma$  는  $0 < \gamma < 1$  로 정의된 학습률(learning rate)이고, 학습자가 신속히 학습하는 데에 기반을 두는 개선된 학습 추정치(learning estimate)  $\Delta Q(s, a)$  는 다음과 같이 표현된다.

**수학식 5**

[0046] 
$$\Delta Q(s, a) = \{R(s, a) + \beta \times \max'_a Q(s', a')\} - Q(s, a)$$

[0047] 여기서  $\beta$  는  $0 < \beta < 1$  로 주어지는 할인된 보상 인자(discounted reward factor)이고,  $\max'_a Q(s', a')$  는 장래의 상태-액션 쌍에 대한 최적의 Q 가치의 추정치를 나타낸다. 종국적으로, Q(s,

a)는 최적의 Q 가치에 수렴하니, 곧  $\lim_{t \rightarrow \infty} Q(s, a) = Q^*(s, a)$  이다. 액션 선택을 위한 한 가지 휴

리스틱(heuristic) 정책은 최대의 측정된 Q 가치를 갖는 액션을 이용하는 것이나, QL은 동적으로 학습 결과를 갱신하는 데에 빈번한 탐색(exploration)을 요구한다. 몇몇 예시적인 구현에서, ReLBT 메커니즘의 QL 기반 모델은 탐색 및 이용(exploration and exploitation) 문제를 다루기 위해 엡실론 탐욕( $\epsilon$ -greedy) 알고리즘을 사용한다(동작(410)). 그러한 ReLBT 메커니즘에서, 제1 확률(가령,  $\epsilon$ )로 탐욕적 액션 정책

$\pi^*(a^*|s) = \operatorname{argmax}_a Q(s, a)$  가 수행되는데, 여기서  $\operatorname{argmax}_a$  는 a에 관해서  $Q(s, a)$ 가

"이용된다"는 것을 나타낸다. 연속적인 이용은 탐욕적인 방식으로 즉각적인 보상을 최대화하는 것으로 이어질 수 있다. 더 자주 이용하는 것이 적당한 대안일 수 있다. 그러나, 예시적인 ReLBT 메커니즘에서, 학습자는 제2

확률(이 값에 전술된 제1 확률이 더해지면 1이 됨)(가령,  $1-\epsilon$ )로 최적 정책  $\pi^*$ 와 상관없이 모든 허용된 액션을 "탐색한다". 이러한 엡실론 탐욕 알고리즘은 시간이 지남에 따라 학습자가 매 액션을 취하면서 Q(s,a)가 수렴되게 한다. 이와 같이, 예시적인 ReLBT 메커니즘에 따르면, 학습자(가령, LTE-LAA 노드)는 그것의 성능을 최적화하기 위해 "이용" 국면(413)에서 동작하고 환경(가령, LTE-LAA 네트워크 환경)에서의 변화를 학습하기 위해 "탐색" 국면(416)에서 동작한다. 이용 국면(413)에서는 전술된 QL 모델(에서 동작(460) 및 동작(462))을 통해 학습된 액션 정책에 따라 ECCA 경쟁 윈도우 값이 조절된다. 탐색 국면(416)에서는 QL과는 별개로, 동작(450)에서, 수학식 2에 제시된 바와 같이, 채널 관측 기반 충돌 확률  $p_{obs}$ 에 기반하여 채널의 유희 여부에 따라 ECCA 경쟁 윈도우 값이 스케일링된다.

[0048] 예시적인 ReLBT 메커니즘의 목적은 LTE-LAA 및 Wi-Fi의 공정한 공정을 최적화하는 것인데, 이는 환경 내의 불필요한 충돌을 감소시킴으로써 달성된다. 따라서, 실시예에서, 채널 관측 기반 충돌 확률  $p_{obs}$ 와 관련하여 임의의 특정 상태에서 수행된 액션의 보상이 표현된다. 예를 들어, 시간 슬롯 t에서 상태  $s_t$ 에서 취해진 액션  $a_t$ 에 의해 주어지는 보상은 다음과 같이 계산된다(동작(460)).

**수학식 6**

[0049] 
$$R_t(s_t, a_t) = 1 - p_{obs}$$

[0050] 위 식은 상태  $s_t$ 에서 액션  $a_t$ 를 수행한 학습자의 만족도를 보여준다고 볼 수 있다. 도 4에 도시된 바와 같이, 학습자는 계속해서 채널 상황을 관측하고 그것의 QL 기반 파라미터를 갱신한다. 학습자는 채널 충돌 확률  $p_{obs}$ 에 기반하여서 그것의 CW를 증가시키거나 감소시킬 뿐이다.

[0051] 도 5는 본 발명의 실시예에 따라 ReLBT 메커니즘이 구현되는 무선 통신 디바이스의 블록도이다. 예를 들어, 무선 통신 디바이스(500)는 네트워크 환경(가령, 도 2의 네트워크 환경(200)) 내에 위치한 LTE-LAA 노드나 Wi-Fi

노드와 같은 디바이스, 또는 이와 통신가능하게 커플링된 다른 디바이스이거나 그러한 디바이스 내에 구현될 수 있다. 도 5를 참조하면, 무선 통신 디바이스(500)는 프로세서(processor)(510), 메모리(memory)(520), 트랜시버(transceiver)(530) 및 안테나(antenna)(540)를 포함하는 것으로 예시된다. 실시예에서, 트랜시버(530)는 프로세서(510)의 제어에 따라 안테나(540)를 통해서 (가령, LTE-LAA 프로토콜에 따라 비인가 대역의) 무선 신호를 송신하는 송신기 및/또는 안테나(540)를 통해서 무선 신호를 수신하는 수신기를 포함할 수 있다. 예를 들어, 무선 통신 디바이스(500)의 송신기는 LTE-LAA 및 Wi-Fi에 의해 공유되는 비인가 대역 내의 채널 상에서의 송신을 위해 구성될 수 있다. 또한, 실시예에서, 프로세서(510)는 전송된 바와 같은 ReLBT 메커니즘을 구현하도록 무선 통신 디바이스(500)를 동작시킬 수 있다. 예를 들어, 메모리(520)에는 컴퓨터 실행가능(또는 프로세서 실행가능) 명령어가 저장될 수 있는데, 이는 프로세서(510)에 의해 실행되는 경우 무선 통신 디바이스(500)로 하여금 (가령, 도 4와 관련하여 기술된 동작을 수행함으로써) 실시예에 따른 ReLBT 메커니즘을 구현하도록 할 수 있다. 몇몇 예시적인 구현에서, 그러한 명령어는 프로세서(510)에 의해 실행되는 경우 프로세서(510)로 하여금 다음과 같은 동작을 수행함으로써 LTE-LAA 및 Wi-Fi의 공존을 위한 그러한 ReLBT 메커니즘을 구현하게 할 수 있다.

[0052] 먼저, 이러한 구현에서, ECCA의 현재 스테이지를 위해 ECCA 경쟁 윈도우 값 CW에 의해 정해진 범위(가령, [0, CW-1])로부터 랜덤 백오프 값이 선택될 수 있다. ECCA는 랜덤 백오프 값으로 백오프 카운터(backoff counter)를 시동하는 것과, 채널의 관측을 위해 ECCA의 현재 스테이지에서의 시간 슬롯(가령, 지속시간  $\sigma$ 의 슬롯) 동안에 채널이 유희인 것으로 감지되면 백오프 카운터를 점감하는 것을 수반하는데, 백오프 카운터가 만료하면(expire) 무선 통신 디바이스(500)의 송신기는 이 채널 상에서 데이터를 송신할 수 있다. 예를 들어, 그러한 채널 관측은 현재의 ECCA 스테이지에서 채널의 유희 여부를 감지하는 것뿐만 아니라 현재의 ECCA 스테이지에서 백오프 카운터가 만료된 후에 수행된 데이터 송신에 대한 피드백(가령, 현재의 TXOP의 HARQ 피드백)을 수신하는 것을 포함할 수 있다.

[0053] 그러면, 채널 관측에 기반하여 의사 충돌 확률(가령,  $p_{obs}$ )이 계산될 수 있다. 이 의사 충돌 확률은 현재의 ECCA 스테이지에서 백오프 카운터가 만료된 후에 수행된 데이터 송신에 대한 NACK(가령, 현재의 TXOP의 HARQ 피드백 내의 NACK)의 개수와 함께, 현재의 ECCA 스테이지에서의 채널 관측 동안에 검출된 분주한 슬롯의 개수 및 현재 선택된 랜덤 백오프 값으로부터 계산될 수 있다(예컨대, 수학적 식 1을 참조하십시오).

[0054] 이제, ECCA의 후속 스테이지를 위해 CW를 조절하는 액션 정책을 학습하기 위한 RL에서 위와 같이 계산된 의사 충돌 확률이 사용될 수 있다. 예를 들어, CW를 조절하는 최적의 액션 정책을 획득하기 위한 RL에는 QL 접근법이 적용될 수 있다. 특히, QL에서 Q 가치를 갱신하는 데에 채널 측정 기반의 의사 충돌 확률이 사용될 수 있는데, 그러한 QL의 보상 함수는 계산된 의사 충돌 확률에 의해 정의될 수 있다(예컨대, 수학적 식 4 내지 수학적 식 6을 참조하십시오).

[0055] 추가적으로, 그러한 후속 ECCA 스테이지를 위한 백오프 값이 랜덤하게 선택될 범위가 새로 정의될 수 있도록, 해당 스테이지를 위해 위와 같이 학습된 액션 정책에 따라(가령, 의사 충돌 확률에 기반하여 갱신된 Q 가치에 따라) CW를 조절하는 이용 국면(가령, 도 4의 블록(413)) 또는 RL과는 별개로 의사 충돌 확률에 기반하여 채널의 유희 여부에 따라 해당 스테이지를 위해 CW를 확대하거나 축소하는(예컨대, 수학적 식 2를 참조하십시오) 탐색 국면(가령, 도 4의 블록(416)) 중 하나에 진입하기로 결정될 수 있다. 앞서 언급된 바와 같이 이용 국면 및 탐색 국면은 각각 제1 확률 및 제2 확률로 진입이 결정될 수 있다(여기서 제1 확률과 제2 확률의 합은 1임). 예컨대, 그러한 결정은 엠실론 탐욕 절차를 통해 수행될 수 있다.

[0056] 예시적인 실시예는 본 문서에 기술된 동작, 기법, 프로세스, 또는 이의 어떤 양상이나 부분이 체현된 컴퓨터 프로그램에 포함하는 컴퓨터 판독가능 저장 매체로서 구현될 수 있다. 이러한 컴퓨터 판독 가능 저장 매체에는 프로그램 명령어, 로컬 데이터 파일, 로컬 데이터 구조 등등이 단독으로 또는 조합되어 포함될 수 있다. 개시된 동작, 기법, 프로세스, 또는 이의 어떤 양상이나 부분을 구현하거나 이용할 수 있는 프로그램은 컴퓨터에 의해 실행될 수 있는 어떤 유형의 (가령, 컴파일형(compiled) 또는 해석형(interpreted)) 프로그래밍 언어, 예컨대, 어셈블리(assembly), 기계어(machine language), 프로시저형(procedural) 언어, 객체지향(object-oriented) 언어 등등으로 구현될 수 있고, 하드웨어 구현과 조합될 수 있다. 용어 "컴퓨터 판독가능 저장 매체"는, 컴퓨팅 장치에 의한 실행을 위한 명령어(실행 시에 컴퓨팅 장치로 하여금 개시된 기법을 수행하게 함)를 저장할 수 있고, 그러한 명령어에 의해 사용되거나 이와 연관된 데이터 구조를 저장할 수 있는 임의의 매체를 포함할 수 있다. 컴퓨터 판독가능 저장 매체의 예는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM, DVD와 같은 광 기록 매체, 플롭티컬 디스크와 같은 자기-광 매체, 그리고 ROM, RAM, 플래시 메모리, 솔리드 스

테이트(solid-state) 메모리와 같은 메모리 디바이스를 포함하되, 이에 한정되지 않는다.

[0057] 성능 평가

[0058] 이제, 예시적인 ReLBT 메커니즘이 기존의 LBT에 비해 LTE-LAA 및 Wi-Fi의 더욱 효율적인 공존을 가능하게 하여 공정성을 개선한다는 것을 보여주는 시뮬레이션 결과가 논의된다. 예시적인 ReLBT 메커니즘의 성능을 평가하기 위해, LTE-LAA 및 Wi-Fi 공존 시나리오(가령, 도 6에 도시된 시나리오(600))가 이용가능한 이벤트 구동형 시뮬레이터 NS3를 사용하였다. 이 시뮬레이션에서, LTE-LAA를 위한 운영자 A(operator-A) 및 Wi-Fi를 위한 운영자 B(operator-B)가 상정되며, 양자 모두 5 GHz 주파수 스펙트럼 내의 동일한 20 MHz 채널을 사용한다. 예시적인 ReLBT의 성능은 지연시간(latency) 및 쓰루풋(throughput)에 대한 두 운영자의 누적 분포 함수(Cumulative Distribution Function: CDF)의 측면에서 기존의 LBT 메커니즘과 비교된다. 도 6은 시뮬레이션 배치 시나리오(600)를 개략적으로 도시한 도면인데, 두 운영자가 4개의 작은 셀 내에 배치된다. 4개의 LTE-LAA eNB 및 4개의 Wi-Fi AP가 자기 위치에 고정된다. 도 6에 도시된 바와 같이, 몇 개의 LTE-LAA UE 및 Wi-Fi STA가 각각 LTE-LAA eNB 및 Wi-Fi 주위에 랜덤하게 분포된다. 표 1에 도시된 바와 같이, 시뮬레이션에서 다음의 두 세트의 밀도가 사용된다: (1) 셀당 UE/STA 개수  $N$ 이 5임(다시 말해, 운영자당 셀 개수가 4이고,  $N$ 이 5이며, 운영자 수가 2이고, 총 UE/STA 개수는  $4 \times 2 \times 5 = 40$ 임); (2) 셀당 UE/STA 개수  $N$ 이 15임(다시 말해, 운영자당 셀 개수가 4이고,  $N$ 이 15이며, 운영자 수가 2이고, 총 UE/STA 개수는  $4 \times 2 \times 15 = 120$ 임).

표 1

파라미터	값
운영자당 셀 개수	4
셀당 디바이스 개수	5, 15
트래픽 모델	UDP 상에서의 FTP
패킷 도달율 ( $\lambda$ )	1.5, 2.5
동작 주파수	5 GHz
채널 대역폭	20 MHz
채널의 물리적 레이트(rate)	MCS 15 (130 Mbps)
데이터 프레임 페이로드(payload)	1000 bytes
$CW_{min}$	LBT: 15, BEB: 15
$CW_{max}$	LBT: 63, BEB: 1023
에너지 검출(Energy Detection: ED) 임계치	LTE-LAA: -72dBm, Wi-Fi: -72dBm
CCA (LTE-LAA), DIFS (Wi-Fi)	60 $\mu$ s, 43 $\mu$ s
슬롯 시간 $\sigma$	LTE-LAA: 9 $\mu$ s, Wi-Fi: 9 $\mu$ s
TXOP (LTE-LAA)	8ms
NACK 피드백 (LTE-LAA)	80%
스케일링 설계 인자 ( $\omega$ )	32

[0059]

[0060]

도 7은 N=5인 경우에 두 개의 상이한 트래픽 도달율  $\lambda=1.5$  및  $\lambda=2.5$ 에 대해 LTE-LAA 및 Wi-Fi 공존의 지연시간 영향을 보여준다. 도시된 바와 같이, 본 발명의 실시예에 따른 ReLBT 메커니즘이 LTE-LAA 및 Wi-Fi 간에 더 공정하다. 더욱이, 도 7(b)에서 볼 수 있듯이, 트래픽 도달율의 증가(즉,  $\lambda=2.5$ )로 공정성 문제가 커지는데, 예시적인 ReLBT 메커니즘은 운영자 간의 공정성을 향상시킬 수 있다. 도 8은 네트워크 내의 총 디바이스 개수가 40에서 120으로 증가한 경우에 트래픽 도달율  $\lambda=1.5$  및  $\lambda=2.5$ 에 대해 LTE-LAA 및 Wi-Fi 공존의 지연시간 영향을 보여준다. 도시된 바와 같이, 셀이나 운영자당 디바이스 개수의 증가는 LTE-LAA 및 Wi-Fi 공존에 더 심각하게 되는데, 예시적인 ReLBT 메커니즘은 LTE-LAA 및 Wi-Fi 운영자 간의 공정한 공존을 개선한다.

[0061]

도 9는 N=5인 경우에 두 개의 상이한 트래픽 도달율  $\lambda=1.5$  및  $\lambda=2.5$ 에 대해 LTE-LAA 및 Wi-Fi 공존의 쓰루풋 영향을 보여주고, 도 10은 N=15인 경우에 두 개의 상이한 트래픽 도달율  $\lambda=1.5$  및  $\lambda=2.5$ 에 대해 LTE-LAA 및 Wi-Fi 공존의 쓰루풋 영향을 보여준다. 도 9에 도시된 바와 같이, 기존의 LBT 메커니즘의 경우에, Wi-Fi STA에는 두드러진 양의 쓰루풋 저하가 있으며, 이는 도 9(b)에서 볼 수 있듯이, 트래픽 도달율이 더 높을수록 더 현저하다. 그러나, 예시적인 ReLBT 메커니즘은 LTE-LAA 노드가 지능적으로 자신의 송신을 수행할 수 있게 하는바,

이는 Wi-Fi STA를 위한 채널 액세스 기회를 증가시키며, 따라서 전체적인 시스템 성능을 개선한다. LTE-LAA 및 Wi-Fi 간의 공정성은 도 10에 도시된 바와 같이, 셀당 디바이스 개수  $N$ 이 증가한 밀집 네트워크 환경에 있어서 더 현저하다. 이는 LTE-LAA UE에 의한 채널 점유 확률 및 시간의 증가 때문이다. 현재 구현된 LTE-LAA LBT 메커니즘은 경쟁 파라미터 갱신을 위해 HARQ 피드백만을 고려한다. 이러한 LTE-LAA LBT 메커니즘으로 인해, 도 10에서 볼 수 있듯이, Wi-Fi 운영자는 현저한 성능 저하에 직면한다. 반면에, 예시적인 ReLBT 메커니즘은 LTE-LAA UE로 하여금 더욱 효율적이고 공정하게 채널 리소스를 액세스할 수 있게 하므로, Wi-Fi의 성능을 개선할 수 있다. 예시적인 ReLBT 메커니즘은 채널 충돌 확률을 사용하여 채널 간섭에 기반하여 CW를 조절하므로, 경쟁자의 수의 증가로 인한 쓰루풋 저하는 도 10에 도시된 바와 같이 기존의 LBT 메커니즘에 비해 예시적인 ReLBT 메커니즘에 영향이 적다. 이 ReLBT 메커니즘의 강화 학습 기반 채널 액세스는 네트워크 내의 LTE-LAA 및 Wi-Fi 디바이스 양자 모두를 위해 더욱 공정한 채널 점유를 가능하게 한다.

[0062] 이상에서 본 발명의 몇몇 실시예가 상세하게 기술되었으나, 이는 제한적이 아니고 예시적인 것으로 간주되어야 한다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 개시된 실시예의 세부사항에 대해 본 발명의 범주에서 벗어나지 않고서 다양한 변경이 행해질 수 있음을 이해할 것이다. 그러므로 본 발명의 범주는 설명된 실시예에 국한되어 정해져서는 안 되며, 후술하는 특허청구범위 및 그 균등물에 의해 정해져야 한다.

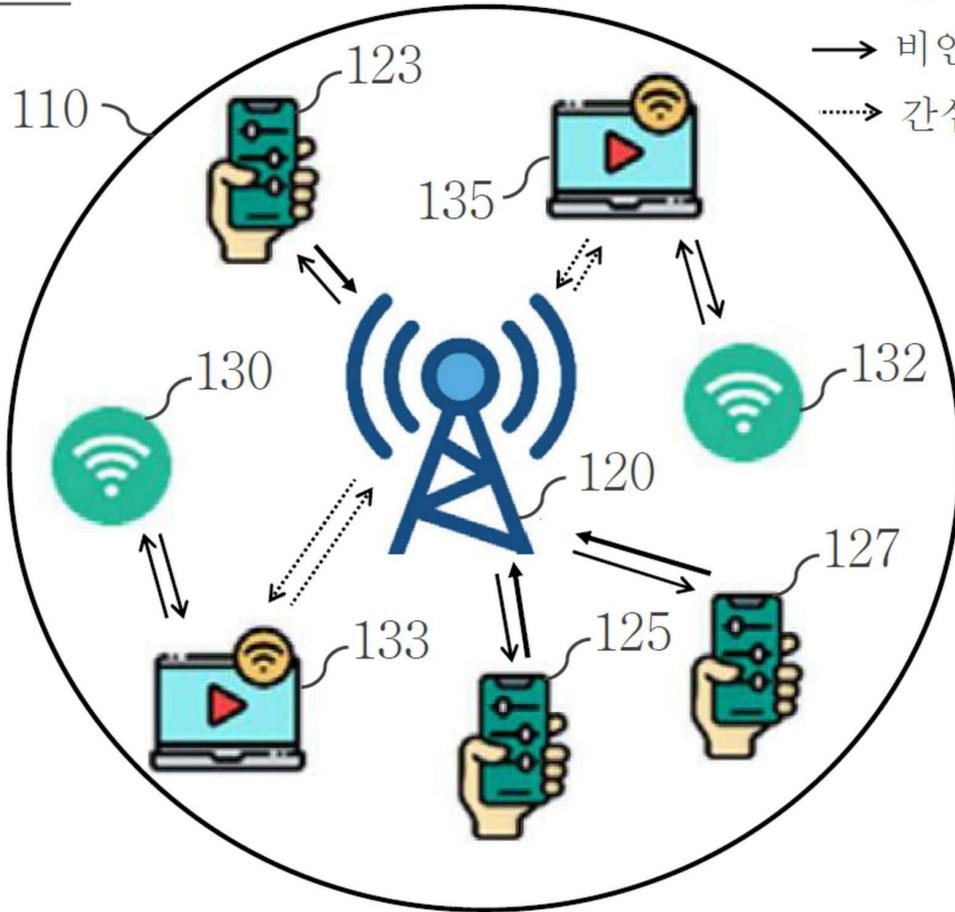
**부호의 설명**

- [0063] 100: 네트워크 환경
- 120: LTE-LAA eNB
- 123, 125, 127: LTE-LAA UE
- 130, 132: Wi-Fi AP
- 133, 135: Wi-Fi STA

도면

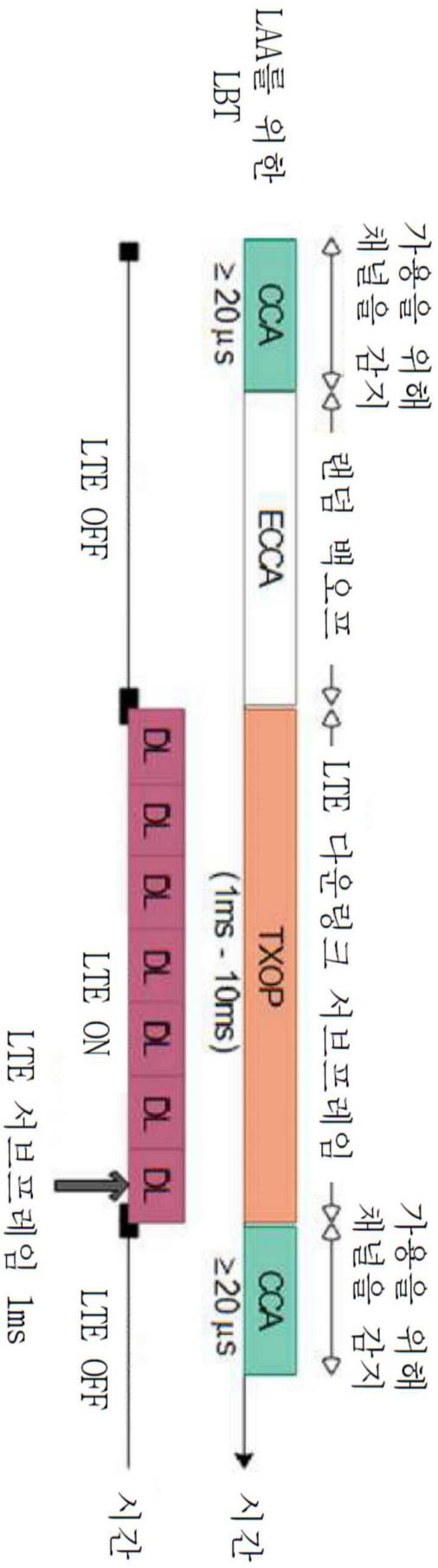
도면1

100

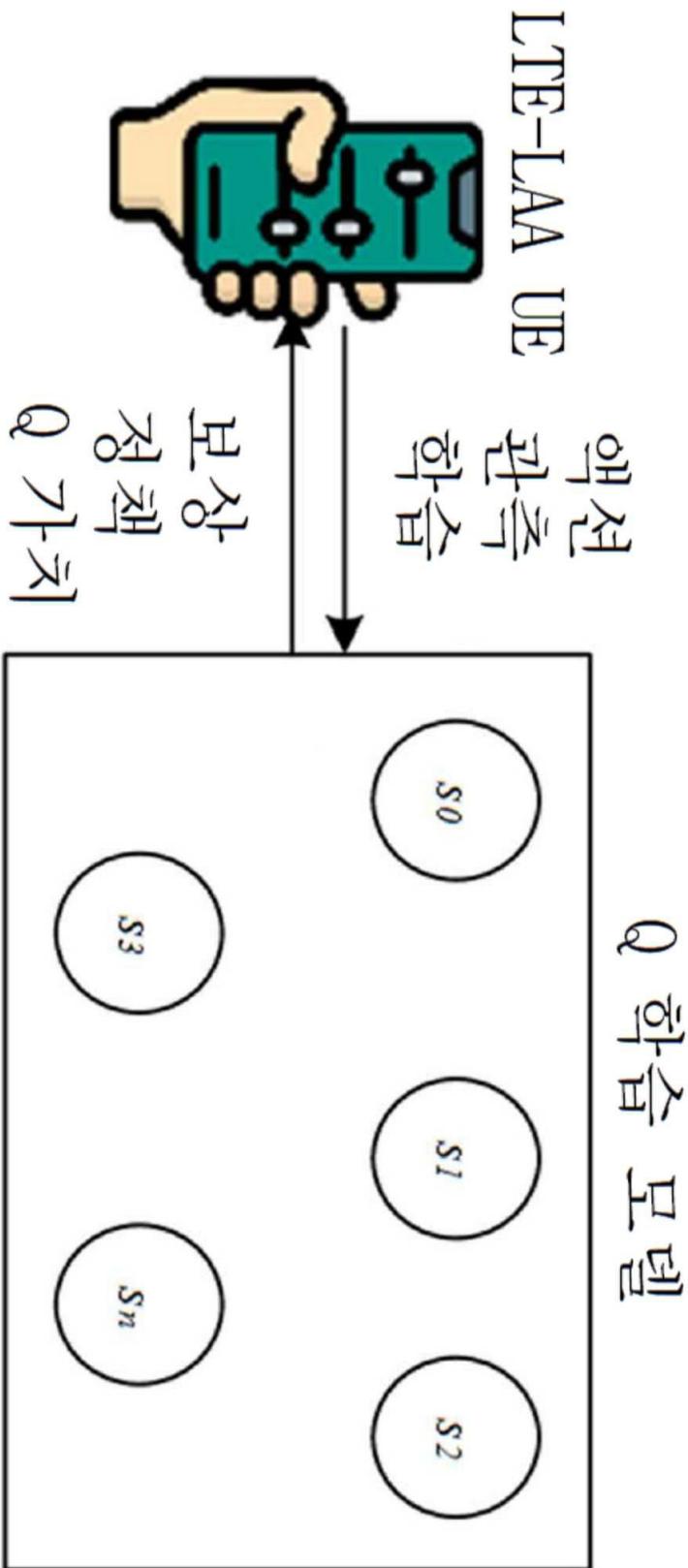


→ 인가 대역 송신  
→ 비인가 대역 송신  
..... 간접

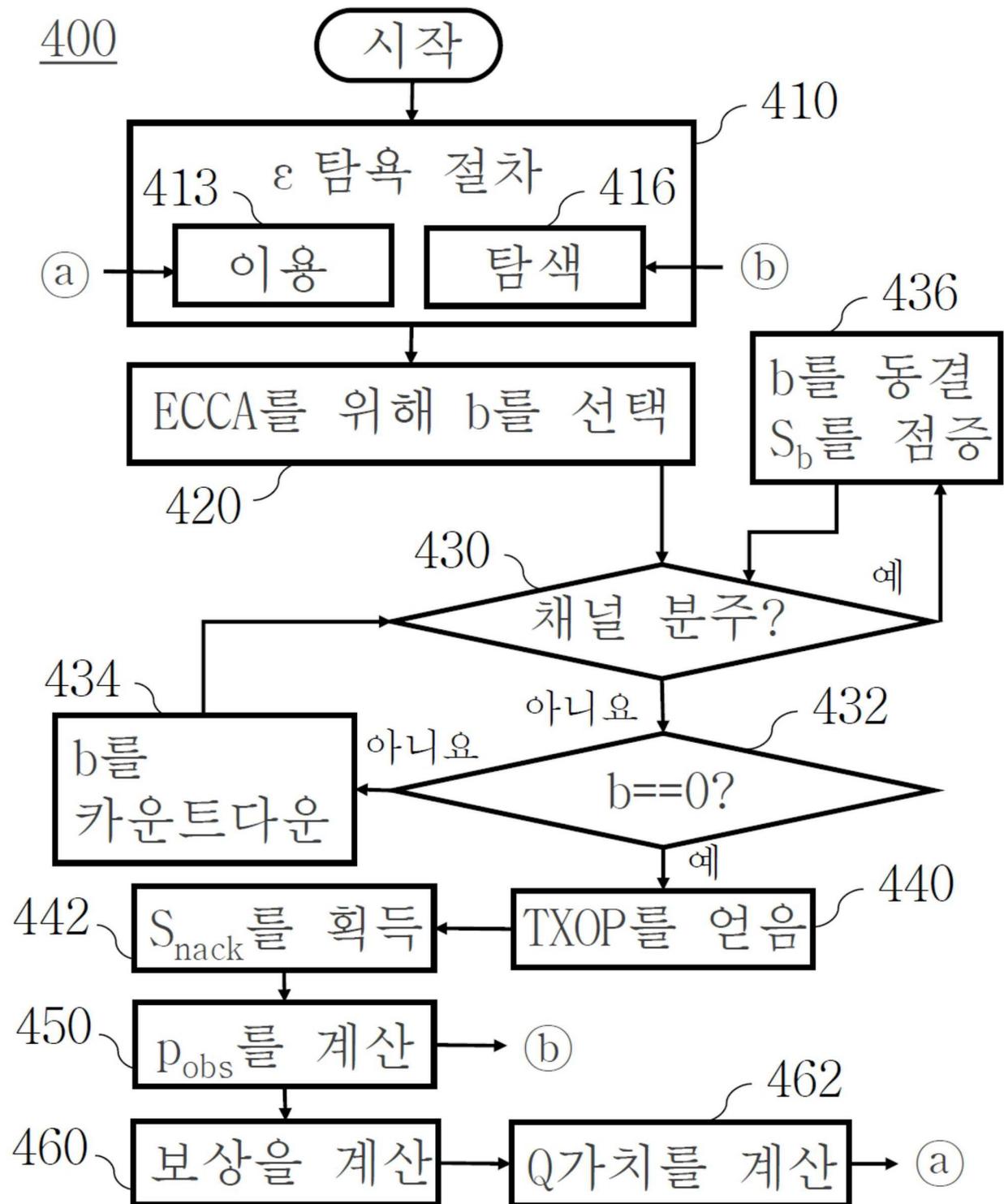
도면2



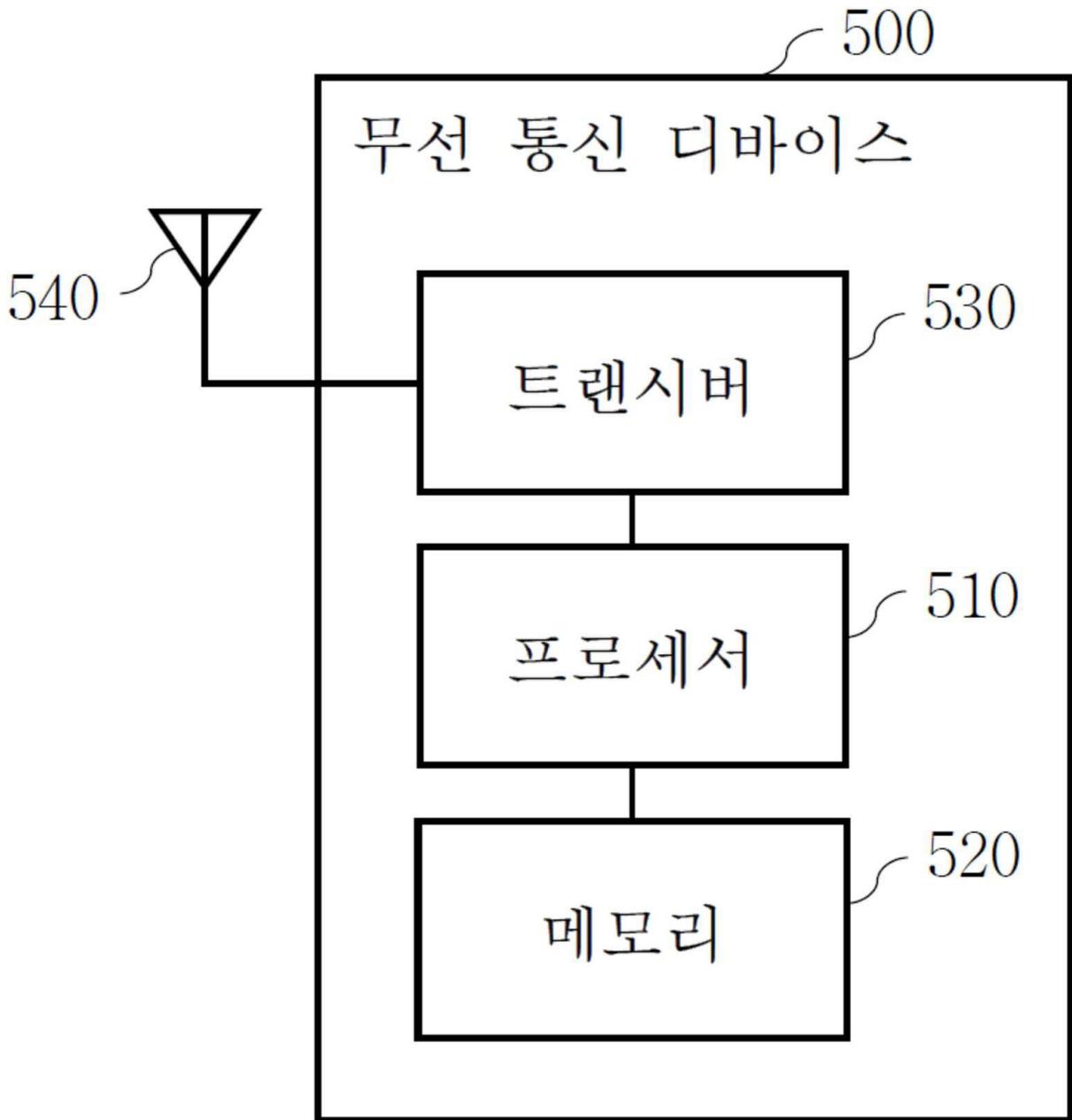
도면3



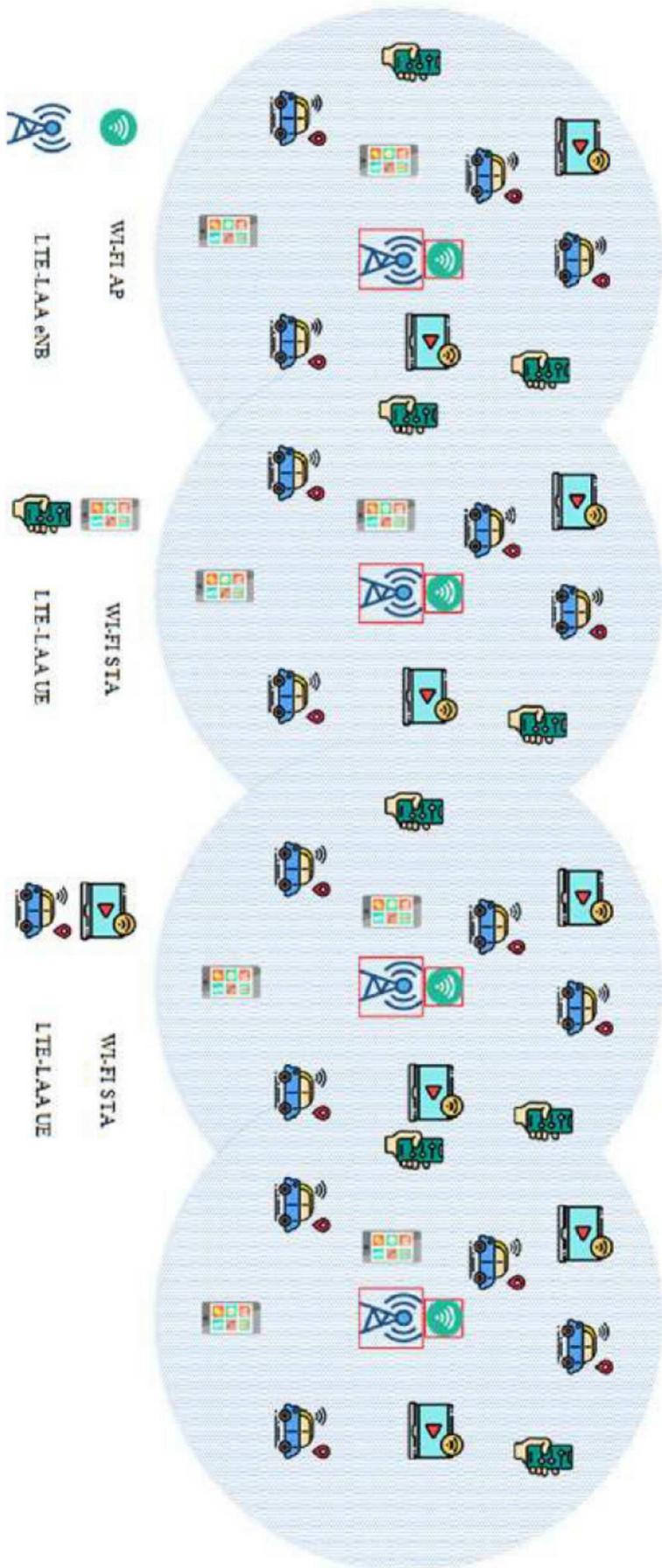
도면4



도면5

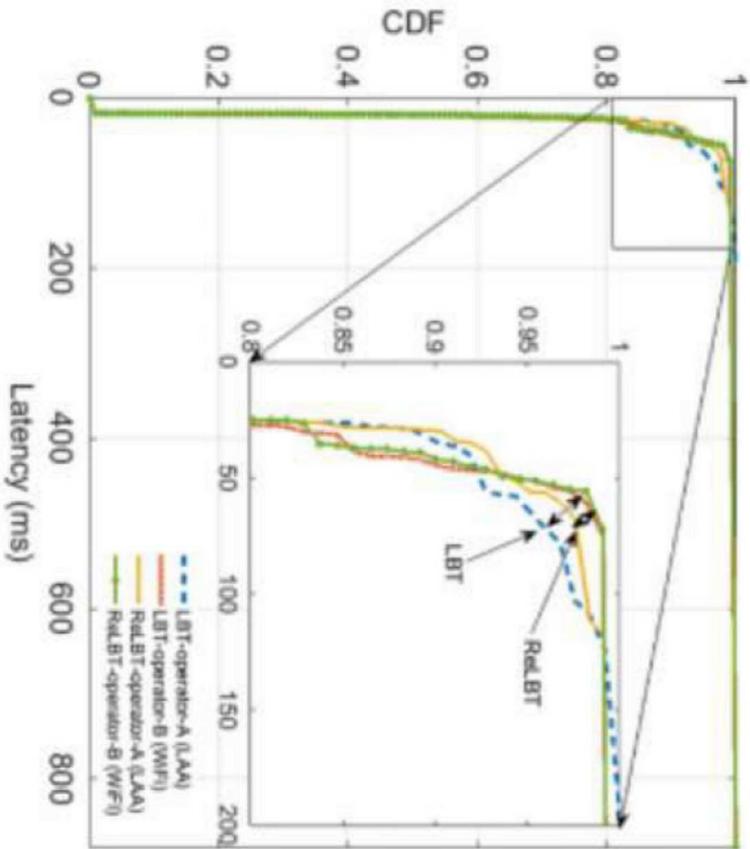


도면6

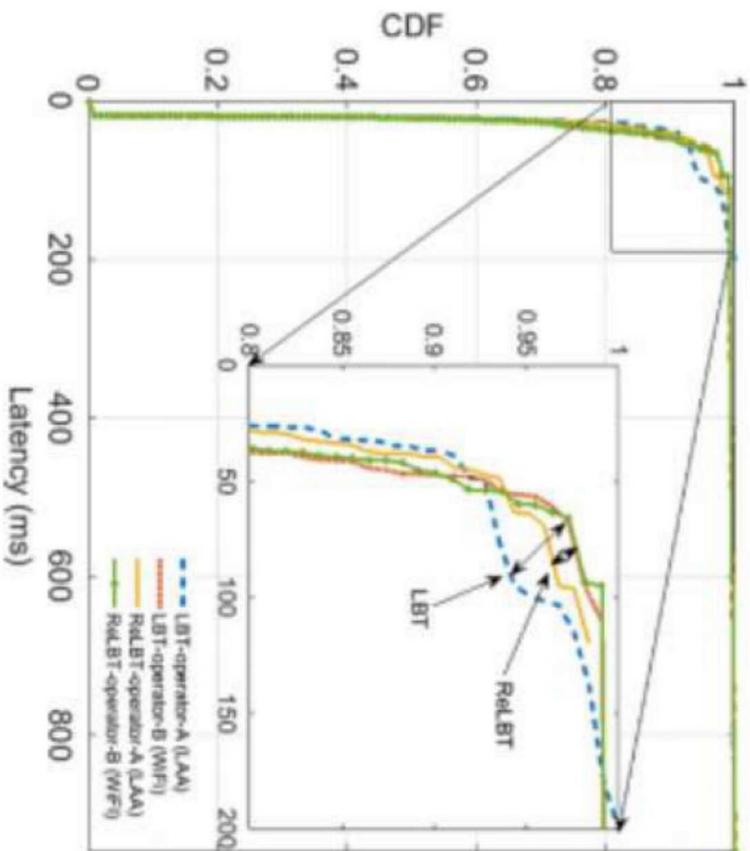


600

도면7

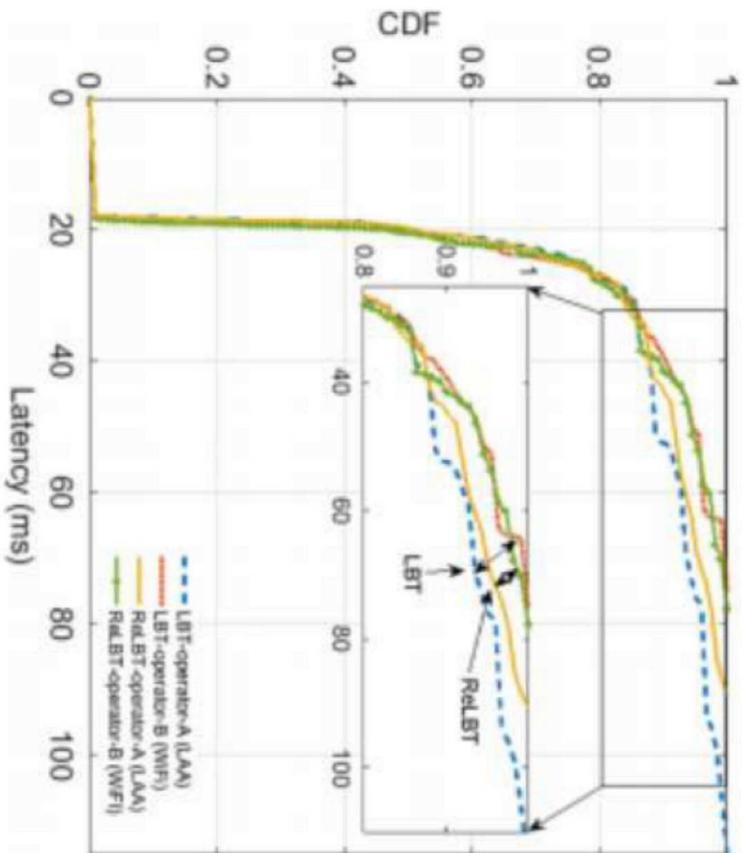


(a)  $N=5$ ,  $\lambda=1.5$

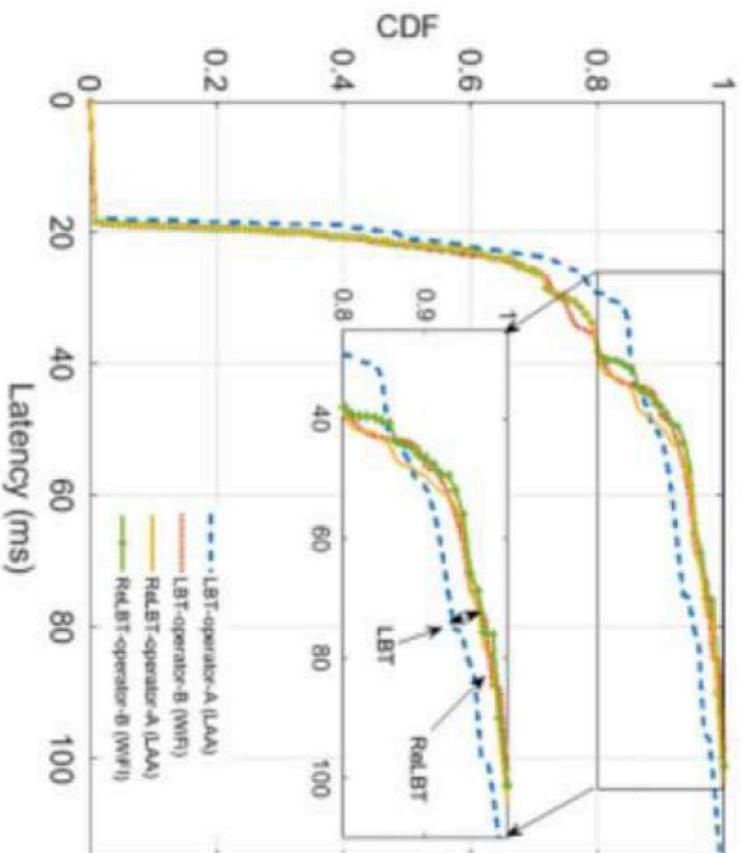


(b)  $N=5$ ,  $\lambda=2.5$

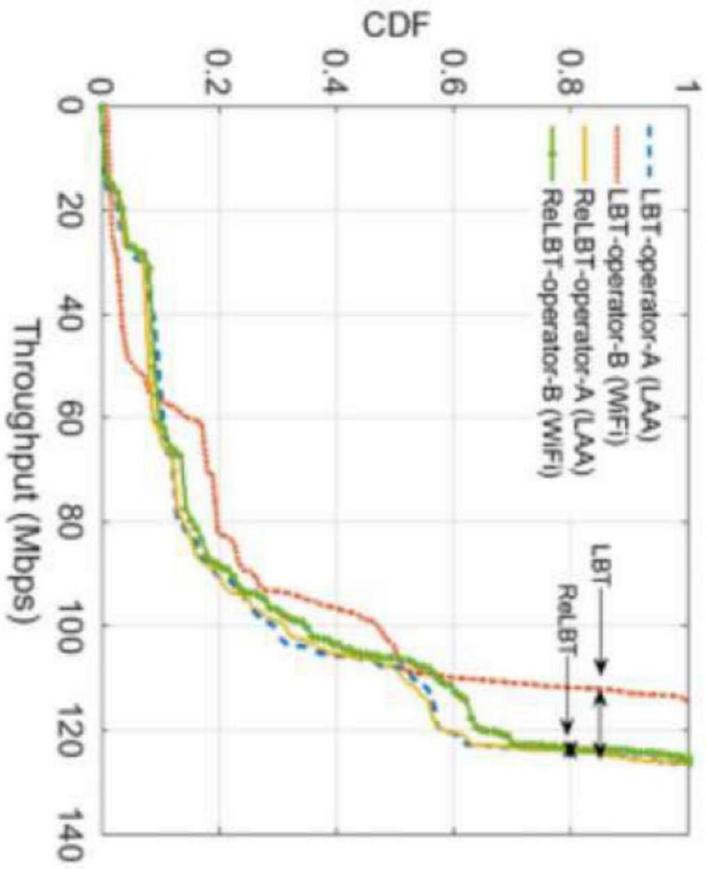
도면8



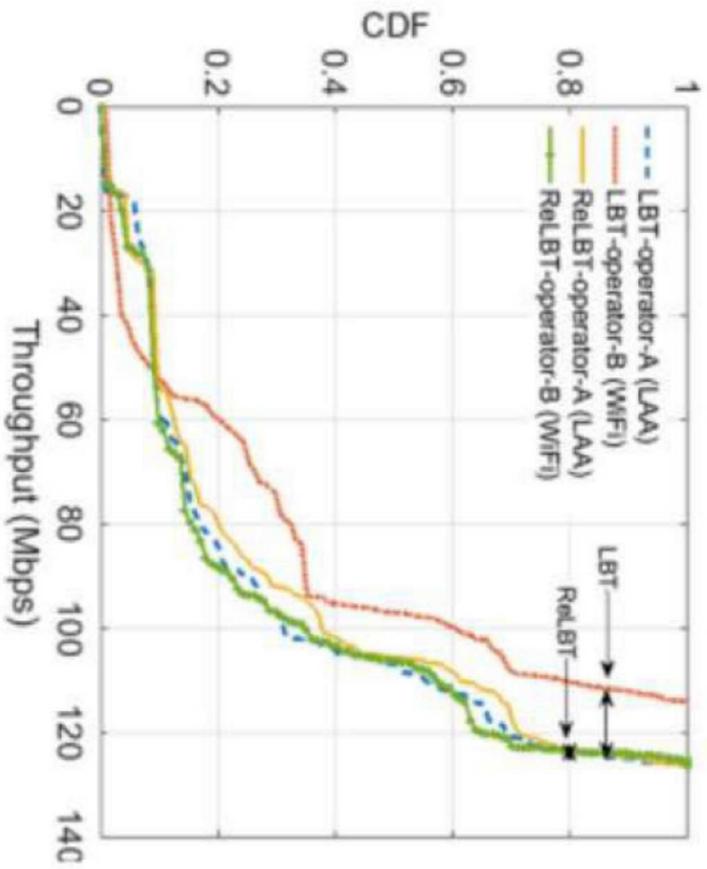
(a)  $N=15$ ,  $\lambda=1.5$



(b)  $N=15$ ,  $\lambda=2.5$



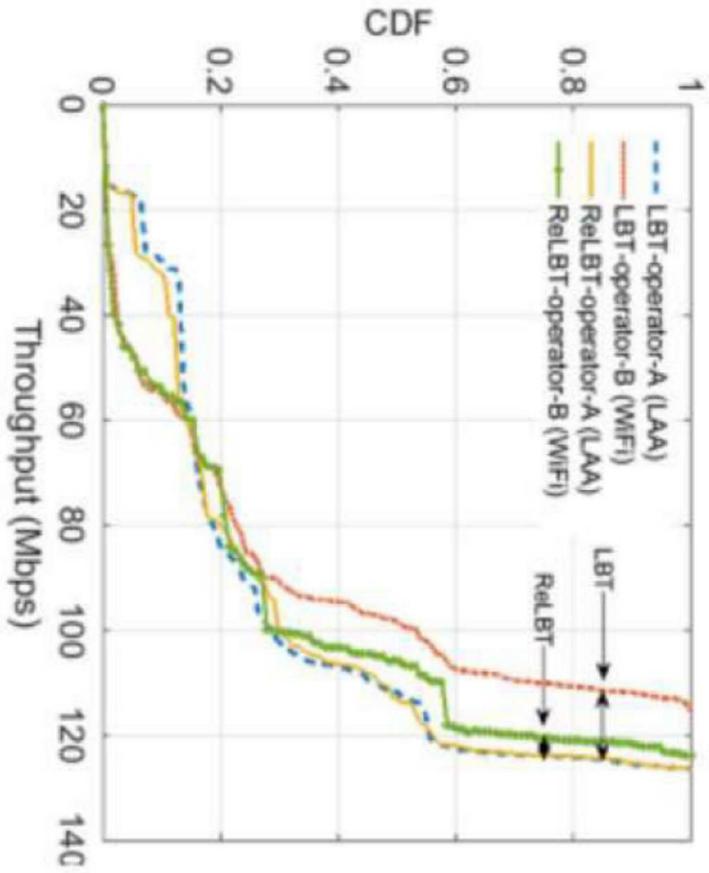
(a)  $N=5$ ,  $\lambda=1.5$



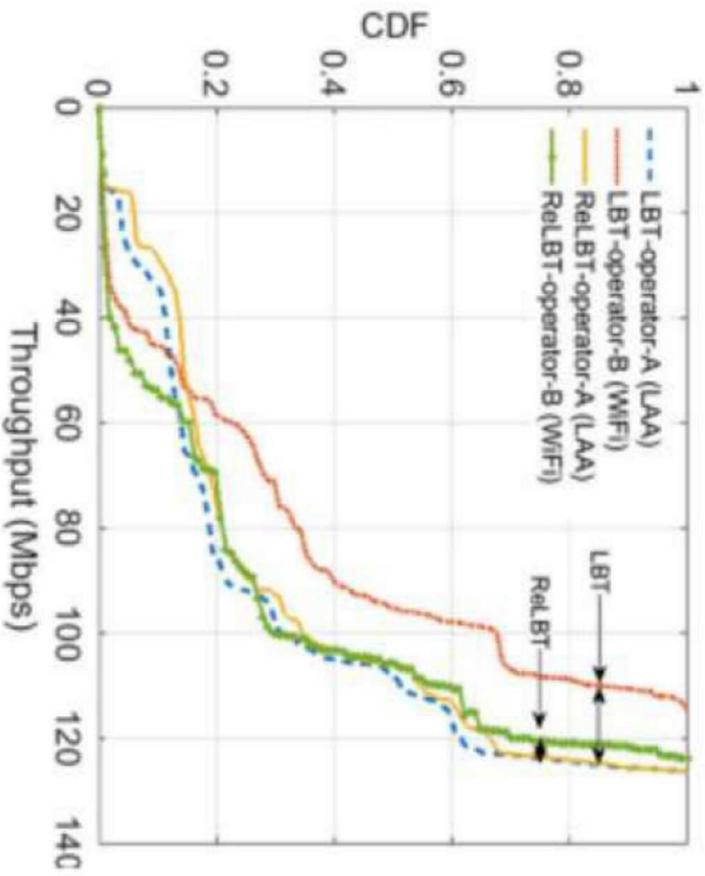
(b)  $N=5$ ,  $\lambda=2.5$

도면9

도면10



(a)  $N=15$ ,  $\lambda=1.5$



(b)  $N=15$ ,  $\lambda=2.5$