한국지능시스템학회 논문지 2012, Vol. 22, No. 1, pp. 12-21
http://dx.doi.org/10.5391/JKIIS.2012.22.1.12

뜨살리스-엔트로피 분석을 통한 무선 랜의 이기적인 노드 탐지 기법

A Study on Detecting Selfish Nodes in Wireless LAN using Tsallis-Entropy Analysis

류병현* · 석승준**+
Byoung-Hyun Ryu and Seung-Joon Seok

* 경남대학교 첨단공학과
** 경남대학교 컴퓨터공학과

요 약

IEEE 802.11 표준 무선 네트워크에서 사용되는 DCF(CSMA/CA) 방식의 MAC 프로토콜은 노드들 사이에서 공평한 채널 접근 확률을 보장하도록 설계되었다. 하지만 최근 급속히 확산되고 있는 무선 환경에서 다른 노드들보다 인위적으로 더 많은 데이터를 전송하는 노드가 존재하는 것이 사실이다. 이들 오동작 노드들은 더 많은 데이터를 보내기 위해서 자신의 MAC 프로토콜 동작을 변형시키거나 다른 노드들의 MAC 동작을 방해한다. 이러한 문제는 이기적(Selfish) 노드 문제라고 정의되어 왔으며, 지금까지의 대부분 연구들에서는 무선 랜 내부의 MAC 프로토콜 동작을 프레임 단위로 분석하여 이기적인 노드를 검색하는 방법을 제안하였으나 모든 종류의 이기적인 노드들을 효과적으로 검출할 수는 없었다. 이러한 단점을 보안하기 위해서 본 논문에서는 통계적 기법 중 하나인 뜨살리스-엔트로피(Tsallis-Entropy)를 사용하여 이기적인 노드 탐색 알고리즘을 제안한다. 뜨살리스-엔트로피는 확률 분포의 밀집도 혹은 분산정도를 효과적으로 나타낼 수 있는 척도이다. 제안한 알고리즘은 무선 랜을 구성하는 AP노드에서 동작하도록 설계되었으며, 무선 노드별로 데이터 간격에 대한 확률 분포를 추출해서 뜨살리스-엔트로피를 계산한 후 임계치와 비교하는 방법으로 이기적인 노드를 검출한다. 논문에서 제안한 이기적 노드 검출 알고리즘의 성능을 평가하기 위하여 다양한 무선 랜 환경(혼잡도, 이기적 노드 동작방법, 임계치)을 고려하여 시뮬레이션을 수행한다. 시뮬레이터는 ns2를 사용하였으며, 실험결과 제안한 방법의 이기적인 노드 검출률이 매우 높음을 알 수 있다.

키워드 : 무선 랜, 이기적인 노드, 탐지 알고리즘, 뜨살리스-엔트로피

Abstract

IEEE 802.11 MAC protocol standard, DCF(CSMA/CA), is originally designed to ensure the fair channel access between mobile nodes sharing the local wireless channel. It has been, however, revealed that some misbehavior nodes transmit more data than other nodes through artificial means in hot spot area spreaded rapidly. The misbehavior nodes may modify the internal process of their MAC protocol or interrupt the MAC procedure of normal nodes to achieve more data transmission. This problem has been referred to as a selfish node problem and almost literatures has proposed methods of analyzing the MAC procedures of all mobile nodes to detect the selfish nodes. However, these kinds of protocol analysis methods is not effective at detecting all kinds of selfish nodes enough. This paper address this problem of detecting selfish node using Tsallis-Entropy which is a kind of statistical method. Tsallis-Entropy is a criteria which can show how much is the density or deviation of a probability distribution. The proposed algorithm which operates at a AP node of wireless LAN extracts the probability distribution of data interval time for each node, then compares the one with a threshold value to detect the selfish nodes. To evaluate the performance of proposed algorithm, simulation experiments are performed in various wireless LAN environments (congestion level, how selfish node behaviors, threshold level) using ns2. The simulation results show that the proposed algorithm achieves higher successful detection rate.

Key Words : Wireless Lan, Selfish Node, Detecting Algorithm, Tsallis-Entropy

1. 서 론

최근 들어 IEEE 802.11 무선 랜(Wireless Lan)을 장착한 스마트 폰의 보급이 늘어남에 따라, 인터넷 사업자들도 공공장소를 중심으로 무선 랜 핫 스팟 영역을 급속히 확대하고 있다. 무선 랜에서는 IEEE 802.11 표

접수일자: 2011년 12월 19일
심사(수정)일자: 2012년 1월 16일
게재확정일자 : 2012년 1월 30일
+ 교신저자
본 논문은 2012학년도 경남대학교 학술연구장려금 지원에 의한 것임

준 MAC(Medium Access Control) 프로토콜인 CSMA/CA(Carrier Sense Multiple Access/Collision Detection) DCF(Distributed Coordinated Function) 매체접근 방식을 기본으로 사용하여 무선 랜 사용자들에게 통계적으로 균등한 데이터 전송 기회를 제공한다. 하지만 최근 이러한 무선 자원의 균등한 사용을 방해하는 오동작(Misbehavior) 노드 문제가 보안 문제로서 대두되어 연구자들의 연구가 이루어지고 있다. 오동작 노드 문제는 세부적으로 무선 랜 내에서 다른 정상 노드들의 데이터 전송을 방해할 목적으로 동작하는 악의적인(Malicious) 노드와 정상 노드보다 더 많은 데이터를 전송하는 것을 목적으로 하는 이기적(Selfish) 노드 문제들로 나눌 수 있다. 본 논문에서는 이 중 후자인 이기적 노드 문제에 관해서 다룬다.

IEEE 802.11 무선 랜 DCF MAC 프로토콜은 충돌윈도우(*CW*)를 사용하여 무선 노드가 데이터 전송 전에 대기해야 하는 시간을 정하도록 한다. 즉, 무선 랜은 데이터 전송을 시도하기 전에 0과 *CW*-1 사이의 랜덤하게 선택한 정수만큼의 타임슬롯(Time-slot)을 대기한 후 프레임을 전송한다. 이 시간을 백오프(Back-off) 지연이라고 한다. 이러한 DCF 동작은 전송할 데이터를 가진 노드들이 서로 다른 백오프 지연시간을 대기함으로서 무선 구간에서 데이터 충돌을 회피할 수 있게 하는 것이다. 하지만 이러한 방법으로도 무선 랜에서 프레임 충돌을 완전히 막을 수는 없다. 다만 이 방법은 여러 단말기들로부터의 전송을 분산시켜서 충돌 확률을 낮추도록 하는데 목적이 있다. 전송된 프레임이 링크에서 충돌되는 경우, 단말기의 DCF 프로토콜은 프레임의 크기에 따라 3회 혹은 6회까지 해당 프레임의 재전송을 시도한다. 충돌이 발생한다는 것은 무선 링크의 혼잡도가 높아졌다는 것을 의미하기 때문에 분산제어 구조를 갖는 DCF 프로토콜에서는 *CW* 값을 충돌시마다 두 배씩 증가시킨다. 이는 증가된 *CW* 값으로부터 얻은 랜덤 타임슬롯 시간을 단말기들이 기다리게 함으로써 동시 전송으로 인한 충돌발생 확률을 줄이기 위함이다. 하지만 *CW* 값이 증가하는 것은 프레임을 무선링크에 내보내기 위해서 단말기가 기다려야 하는 백오프 지연시간이 길어지는 것을 의미한다. 또한 DCF는 백오프 지연시간을 랜덤하게 정하기 때문에 *CW* 값이 커지면 백오프 지연시간의 변화폭도 당연히 커지게 된다. 이와 같이 무선 랜 단말기들은 상호 독립적인 랜덤 시간을 사용함으로서 통계적으로 균등한 데이터 전송을 얻을 수 있게 된다. 그림 1은 IEEE 802.11 DCF MAC 동작 메커니즘을 그린 그림이다[1].

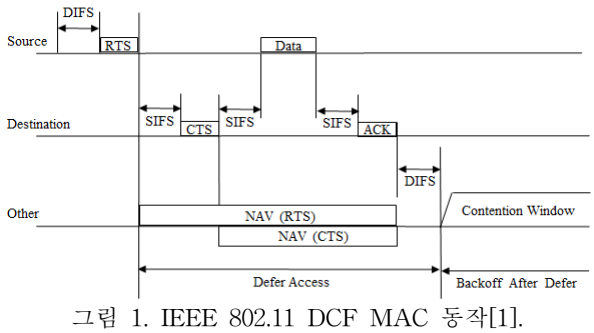

그림 1. IEEE 802.11 DCF MAC 동작[1].
Fig. 1. IEEE 802.11 DCF MAC operation[1].

이기적 노드는 자신의 MAC 프로토콜의 다양한 동작 파라미터를 임의적으로 조작하여 표준 DCF 방식을 따르는 정상 노드들과 균등하지 못한 경쟁을 해서 무선 환경에서 먼저 채널을 점유하여 결국 더 높은 전송률을 획득하는 것이 목적이다. 이러한 이기적인 노드는 무선 네트워크를 관리 관점에서 보면 무선 채널이 혼잡하지 않는 경우에는 큰 문제가 되지 않지만 최근과 같이 공중 무선 랜이 혼잡한 경우에는 이기적인 노드의 존재는 심각한 문제가 된다.

본 논문에서는 무선 랜을 구성하는 노드들 사이에서 이기적 노드를 검출하는 효과적인 알고리즘을 제안한다. 기존의 많은 연구에는 주로 MAC 프로토콜 동작을 추적하여 정상적이지 못한 이기적 노드를 검출하는 방법을 제안하였다. 하지만 기존 논문들에서 제시하는 대부분의 동작기반 방법들은 실제 가능한 다양한 이기적 노드 동작들을 고려하지 못하고 있기 때문에 최근 들어 몇몇 연구에서는 무선 랜 혹은 노드의 성능을 측정하여 통계적 기법으로 이기적 노드를 탐지하는 방안을 제안하고 있다. 본 논문에서는 기존의 방법과는 다르게 수집된 데이터를 통계적으로 분석하여 이기적인 노드와 정상적인 노드를 탐지하는 알고리즘을 제안한다. 이기적인 노드 탐지 알고리즘은 각 노드로부터의 데이터 전송 간격에 대한 확률 분포를 추출하고 뜨살리스-엔트로피(Tsallis-Entropy)를 계산하여 이기적 노드를 검출하도록 한다. 또한 제안하는 알고리즘은 이동 평균 모델을 활용하여 동작하는 무선 랜에 대한 시계열 분석을 시도한다.

2. 기존 연구

이기적인 노드는 무선 환경에서 정상 노드보다 좀 더 빠르게 전송하기 위해서 자신의 MAC 파라미터를 불법적으로 수정을 하거나 강제적인 충돌함으로써 신의 전송량을 높여서 데이터를 전송한다. 이기적인 노드가 조작하는 주요 MAC 파라미터로는 DIFS, NAV, CW가 있다. 이밖에도 이기적인 노드는 정상 노드의 데이터에 대해 인위적으로 충돌을 발생시켜 해당 정상 노드의 백오프 동작을 유발시키는 방법을 사용할 수도 있다. 이러한 이기적 노드 문제를 해결하기 위해서는 무선 랜 내에서 동작하는 이기적 노드를 정확하게 검출할 수 있는 방법이 먼저 연구되어야 한다. 지금까지 연구된 이기적 노드 검출 방식에는 프로토콜 기반 검출 방식과 통계 기반 검출 방식으로 나눌 수 있다.

2.1 프로토콜 기반 이기적인 노드 검출 기법

정상 노드와 이기적인 노드를 구별하는 프로토콜 기반 이기적인 노드 검출 기법은 특정 파라미터를 수정한 이기적인 노드를 검출 할 수 있는 장점이 있고 다양한 이기적인 노드를 효과적으로 검출할 수가 없다는 단점이 있다. 프로토콜 기반 이기적인 노드 검출 기법으로 다양한 기법들이 제안되어 있지만, 여기에서는 대표적인 두 가지 기법에 대하여 설명한다. 첫 번째 프로토콜 기반 이기적인 노드 검출 기법은 백오프 시간을 비교하여 이기적인 노드를 검출하는 기법이다[2]. 이기적인 노드는 백오프 시간을 정상 노드보다 반 이상 줄였기 때

문에 무선에서 충돌이 일어났을 경우에 정상 노드보다 채널 전송 간격이 짧은 특징을 가진다. 그러므로 실제 백오프 구간을 측정해서 기존의 백오프 값과 비교하여 이기적인 노드를 탐지할 수 있다. 그림 2는 실제 백오프 구간을 측정하여 이기적인 노드를 검출하기 위한 구조를 그린 그림이다.

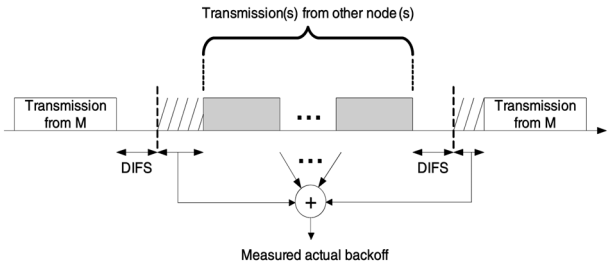

그림 2. 백오프 시간 비교를 통한 이기적인 노드 검출 기법[2].

Fig. 2. Selfish node detection method through comparing back-off interval[2].

두 번째 기법은 수신측에서 수신되는 프레임의 절차를 확인하는 기법이다. 정상적으로 DCF 프로토콜을 따르는 프레임은 RTS, CTS, DATA, ACK 순으로 수신측에 순서대로 수신된다. 하지만 이기적인 노드는 많은 전송량을 원하기 때문에 강제적으로 정상 노드의 RTS 신호에 맞춰서 충돌을 일으켜 수신측에서 RTS를 수신하지 못하도록 하는데, 이 기법은 수신측에서 정상적인 프레임의 절차 유무를 판단하여 이기적인 노드를 판별한다[2]. 또한, MAC 계층에서 이기적인 노드를 검출하는 기법 외에 PHY 계층에서 존재하는 파라미터를 수정하여 이기적인 노드를 생성하는 방법이 있다. 이 방법은 이기적인 노드가 자신의 CCA(Clear Channel Assessment) 임계치를 올려 정상 노드가 전송하는 무선 데이터의 전력량을 측정하고 자신의 임계치보다 낮으면 전송하지 않다고 판단하여 자신이 전송하는 방법이다. 이러한 노드의 해결방법은 AP에서 낮은 전력 프로브 메시지를 전송범위에 존재하는 모든 노드에게 전송하여 응답을 하지 않는 노드를 이기적인 노드라고 판단하는 방법이다[3].

2.2 통계 기반 이기적인 노드 검출 기법

통계 기반 이기적인 노드 검출 기법은 무선 채널에서 수신자 역할인 AP가 일정 구간 동안에 주기적인 모니터링을 반복하여 성능을 측정하는 방식이다. 이 기법은 다양한 이기적인 노드를 검출할 수 있다는 장점이 있으며 본 절에서는 대표적인 두 가지 기법에 대하여 설명한다. 첫 번째 통계 기반 검출 기법은 K-S(Kolmogorov-Smirnov) 검정법을 사용한 이기적인 노드 검출 기법이다. 이 기법은 확률 분포들을 상호 비교 하는 경우에 사용되고 이기적인 노드를 검출하는 데 K-S 검정법을 사용한 두 가지 사례가 있다. 첫 번째는 노드들의 확률분포가 균등 분포(Uniform Distribution)를 따르는 가를 판단하여 이기적인 노드를 검출하는 방법이다. 이 방법에서는 전송에 성공한 프레임의 시간 간격을 수집한다. 만약 재전송을 한 프레임의 시간과 재전송을 하지 않는 프레임의 시간을 둘 다 수집하였을 때, 무선에서 충돌이 일어났을 경우 백 오프 메커니즘을 사용하였기 때문에 이기적인 노드를 탐지하는 데 불필요한 수집 대상으로 판단하여 재전송을 한 프레임의 시간 간격을 제외하는 방법을 사용한다. 그래서 전송에 성공한 연속적인 두 개의 프레임의 간격을 측정해 이기적인 노드를 탐지한다. 전송에 성공한 프레임의 간격을 수집해서 균등 분포로 변환하였을 때 일정하게 나왔다면 이 파라미터는 균등 분포를 따른다. 그러므로 각 노드별로 균등 분포를 따르고 연속적이며 전송에 성공한 프레임의 간격을 측정하여 K-S 검정법을 통해 이기적인 노드를 탐지한다[4]. 두 번째는 노드별 전송 간격을 상호 비교해서 분포의 크기를 비교한 방법이다. 이 방법에서 이기적인 노드는 전송간격이 정상 노드보다 짧기 때문에 전송간격의 확률분포가 정상 노드보다 밀집되어 있다는 점을 이용해 가장 먼저 각 노드별 프레임전송 간격을 추출한 후 서로 다른 두 노드의 프레임 전송 간격으로 확률 분포로 만들고 K-S 검정법을 사용하여 두 노드의 확률 분포를 비교한다. 결과적으로 모든 노드들의 확률 분포가 일치하는 그룹과 일치하지 않는 그룹으로 분류한 후에 탐욕 노드 그룹을 분류한다. 그룹 안의 전송량의 분산 값을 임계치와 비교하여 탐욕 노드로 분류하여 탐욕 노드들의 접근을 차단하는 기법이다[5].

두 번째 통계 기반 검출 기법은 전송 빈도 비교 탐지법이다. AP는 이기적인 노드를 탐지할 때 모든 노드들로부터 프레임의 수신 빈도를 측정한다. 특정 수신 구간을 지정해 그 시간 동안에 데이터를 전송한 모든 노드의 프레임의 횟수를 관찰한다. 그림 3은 각 노드별로 전송수를 측정하여 비교한 그래프이다. 노드별로 전송 빈도를 비교하여 월등히 전송 빈도가 높은 노드는 이기적인 노드로, 전송 빈도가 낮은 노드는 정상 노드로 판별하는 방법이다[6].

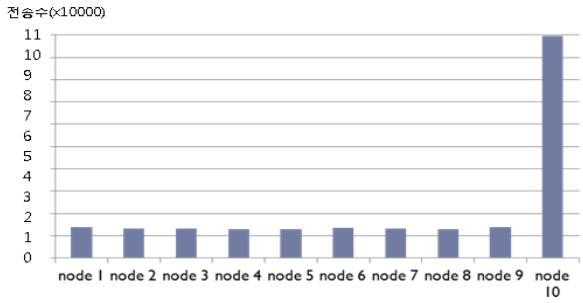

그림 3. 데이터 전송 빈도 비교.

Fig. 3. Comparing the frequencies of data transmission.

3. 엔트로피 정의 및 활용방법

샤논은 무질서도의 척도로서 엔트로피라고 개념을 제안하였다[7]. 엔트로피는 생활과 수학, 일상적인 부분에서 나타나는 신호와 정보에 대한 불확실성을 수치화하는데 사용되고 식 (1)과 같이 엔트로피 H를 정의하

였다.

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log_e p(x_i) \quad (1)$$

여기서, X는 확률변수이고 $p(x_i)$는 확률변수 X의 값이 $x_1 \sim x_n$값을 가질 확률이다.

통신 네트워크에서 최근 엔트로피 활용 사례로는 DDOS(Distributed Denial of Service), Network Scan, Port Scan 탐지 등 네트워크 보안 감시 분야에서 이상 데이터 흐름을 순간적으로 탐지하기 위해서 엔트로피 분석법을 활용하고 있다. 본 논문에서는 앞서 언급한 무선 랜의 이기적 노드를 탐지하기 위하여 데이터 전송 간격에 대한 엔트로피 분석법을 활용한다. 특히, 샤논-엔트로피보다 더 세밀한 분야에 사용되는 뜨살리스-엔트로피를 활용한다. 뜨살리스-엔트로피는 기존의 샤논 엔트로피보다 정밀한 무질서도 양의 분석이 가능한 기법이다. 무선 네트워크에서는 이기적인 노드를 탐지하는 방식이 아주 미세한 파라미터를 표본의 대상으로 지정하기 때문에 기존의 샤논-엔트로피보다 뜨살리스-엔트로피를 사용하면 더 나은 분석결과를 얻을 수 있다. 뜨살리스-엔트로피는 식 (2)과 같이 정의된다.

$$H_q(X) = \frac{1}{q-1}(1 - \sum_{i=1}^{n}(p(x_i))^q \quad (2)$$

여기서, X는 확률변수이며 $p(x_i)$는 확률변수 X의 값이 $x_1 \sim x_n$값을 가질 확률(밀도)이고 q는 1에 가까우며 실제 네트워크 환경에 적용할 경우 변할 수 있는 값이다. 확률 분포에서 앞부분의 확률이 높을 경우 $q > 1$, 뒷부분의 확률이 크면 $q < 1$을 사용한다. 그림 4는 9개의 정상노드와 1개의 이기적인 노드가 혼재하는 무선 랜 환경에서 실시된 시뮬레이션에서 각 노드가 전송하는 데이터 간격에 대한 분포를 얻어 엔트로피 분석을 수행한 결과이다.

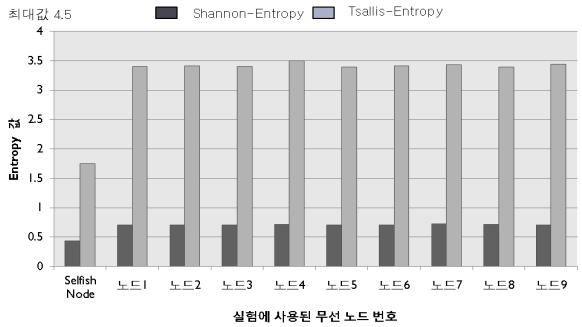

그림 4. 노드별 샤논-엔트로피와 뜨살리스-엔트로피 비교.

Fig. 4. Comparing shannon-entropy and tsallis-entropy values for each node.

실험결과 동일한 전송 간격 분포들에 대해서 샤논-엔트로피보다 뜨살리스-엔트로피 결과 값들이 정상노드와 이기적 노드 사이의 차별화를 잘 표현하고 있다. 이와 같은 점을 활용하여 본 논문에서는 뜨살리스-엔트로피 기반 이기적 노드 검출 알고리즘을 제안한다.

4. 뜨살리스-엔트로피를 이용한 이기적인 노드 검출 알고리즘

본 논문에서는 무선 랜에서 정상 노드보다 인위적으로 데이터 전송률을 높이려고 하는 이기적 노드를 검출을 위해 AP에서 동작하는 알고리즘을 제안한다. 인프라스트럭처 형태의 무선 랜에서는 모든 데이터 패킷이 AP에 의해 중계되기 때문에 AP는 제안하는 알고리즘을 수행하기 위해 적절한 장치가 될 수 있다. 제안하는 알고리즘은 이기적 노드의 전송 패킷 간격이 확률적으로 짧을 것이라는 가설을 기반으로 설계된다. 이는 이기적 노드의 전송률이 더 높기 때문이다. 또한, 이기적 노드는 인위적인 방법으로 전송 간격을 줄이려는 방법을 사용하기 때문에 전송 간격에 대한 확률 분포가 정상 노드에 비해서 상대적으로 밀집될 것이다. 즉, 이기적 노드의 전송 간격에 대한 엔트로피 값이 정상 노드에 비해 작을 것이다.

본 논문에서는 이기적인 노드를 탐지하기 위해 두 가지 파라미터 즉 확률 변수에 대해 엔트로피 분석을 수행한다. 첫 번째 확률 변수는 각 무선 노드로부터 전송되는 데이터의 시간 간격(Between Packet Arrival Time : BPAT)이고 두 번째 확률 변수는 일정 시간 윈도우 동안에 데이터 패킷의 간격 중 최댓값(Window-Maximum between packet arrival time : WM)을 추출한 것이다. 사전 실험결과 일반적인 혼잡 상황에서는 전송 간격에 대한 엔트로피 분석만으로도 이기적 노드를 효과적으로 검출할 수 있다. 하지만, 극심한 혼잡 상황 혹은 혼잡하지 않는 무선 랜에서는 최대 시간 간격에 대한 엔트로피 분석이 효과적인 결과를 얻을 수 있다. 두 경우들을 모두 고려할 수 있는 일반화된 알고리즘을 제안하기 위해서 논문에서는 두 엔트로피를 비교하여 더 작은 값을 임계치와 비교하여 이기적 노드를 탐지한다. 그림 5는 논문에서 제안한 이기적인 노드 탐지 알고리즘을 3단계로 나눈 그림이다. 동작중인 무선 랜의 상태를 지속적으로 감시하기 위하여 논문에서는 그림 3의 알고리즘은 일정 시간 간격으로 반복적으로 수행되도록 한다. 또한 무선 랜에 대핸 시계열 분석을 위해 계산된 엔트로피 값에 대한 이동 평균을 구한 후 임계치와 비교한다.

알고리즘의 첫 번째 단계에서는 각 노드별로 엔트로피를 만들기 위해서 데이터를 수집하고 분류한다. 무선 노드들로부터 수신되는 프레임을 AP에서 추출하고 들어오는 프레임의 헤더 정보에 따라서 소스 노드별로 분류한 후 연속된 프레임간 시간 간격을 구한다. 또한, 시간 윈도우 구간마다 최댓값을 추출한다. 두 번째 단계에서는 각 주기마다 분류된 두 가지 파라미터(확률 변수)로부터 뜨살리스-엔트로피를 계산한다. 즉, 추출된 두 확률 분포로부터 뜨살리스-엔트로피를 각각 계산한

다. 그 다음 세 번째 단계에서 뜨살리스-엔트로피 값에 대해서 시계열 이동 평균(Moving Average), 정규화, 그리고 임계치와 비교를 통해 이기적 노드를 검출한다. 상세한 알고리즘 단계별 동작은 아래에 기술한다.

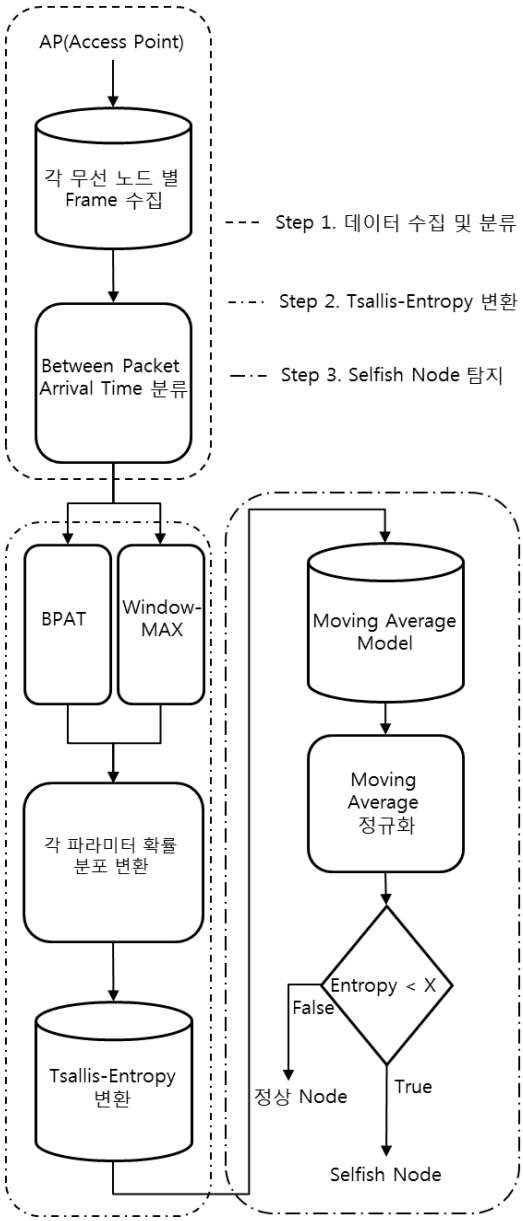

그림 5. 이기적인 노드 탐지 알고리즘 순서도.
Fig. 5. Flow chart of selfish node detection algorithm.

4.1 데이터 수집 및 분류

본 절에서는 이기적 노드 알고리즘에서 사용될 확률 분포를 추출하기 위하여 두 가지 파라미터에 대해서 데이터를 수집하고 분류하는 작업에 대하여 설명한다. 데이터 수집 및 분류 단계에서는 먼저 AP에 수신되는 프레임을 송신 노드별로 분류한 후 각 노드별 두 가지 파라미터를 추출하는 작업을 수행한다. 이 작업은 송신 노드별 확률 분포를 추출한 후 엔트로피 값을 비교하여 무선 노드가 이기적 노드 유무를 검사하는 알고리즘의 첫 단계 작업이 된다.

첫 단계에서 AP는 수신되는 모든 프레임 중에서 RTS 프레임을 추출한다. 무선 랜에서는 데이터 전달을 위해 무선 노드와 AP 간 RTS-CTS-DATA-ACK 프레임이 순서대로 교환되기 때문에, 제안하는 알고리즘에서는 AP에 수신되는 연속적인 RTS 프레임 간격을 측정하여 필요한 두 가지 확률 변수 값을 얻는다. 논문에서 사용하는 두 확률 변수는 앞서 언급한 것과 같이 단순한 데이터 간격과 시간 윈도우 구간에서 최대 데이터 간격이다. 그림 6은 두 가지 확률 변수의 예를 나타낸 그래프이다.

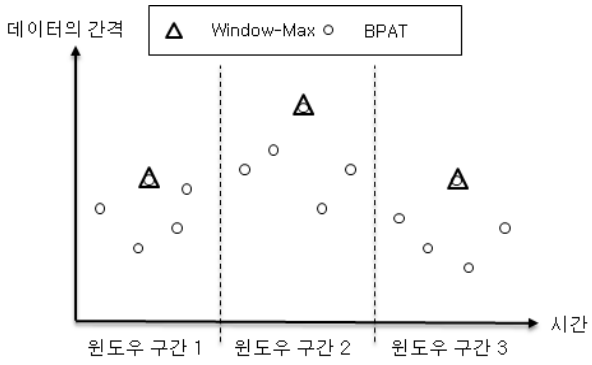

그림 6. 확률 변수 예.
Fig. 6. Example of random variables.

4.2. 뜨살리스-엔트로피 변환

앞서 언급하였듯이 본 논문에서는 효과적으로 무선 랜의 이기적인 노드를 탐지하기 위해서 확률적인 접근 방식인 뜨살리스-엔트로피 분석방법을 사용한다. 뜨살리스-엔트로피(식 (2))를 적용하기 위해서는 p(x)의 확률 변수가 정의되어야 한다. 이를 위해 알고리즘의 2단계에서는 앞서 1단계에서 추출된 두 확률 변수에 대해서 확률 분포를 계산하는 작업을 먼저 수행한다. 두 확률 변수에 대해 확률 분포를 계산하는 식은 다음과 같다. 앞서 언급한 두 확률 변수는 연속적인 값(continuous value)을 갖는 확률 변수 이다. 따라서 본 논문에서는 확률 변수 BPAT와 WM에 대한 엔트로피 분석을 위해 손쉽게 확률 분포를 계산하기 위해서 다음과 같은 식을 사용한다.

$$p(x_i) = R_i \times \frac{1}{\sum_{i=1}^{n} R_i} \tag{3}$$

여기서, R_i는 각 파라미터($BPAT$, WM) 값을 n개의 구간으로 나눈 후 각 구간별 발생 빈도를 측정한 값이다. $\sum_{i=1}^{n} R_i$는 총 발생빈도이다. 구간의 크기가 작을수록 분석은 알고리즘의 정확도는 높아지는 대신 계산의 복잡성 또한 높아지는 단점이 있다.

4.3 이기적인 노드 탐지

알고리즘의 마지막 단계는 시계열 분석을 통한 이기적 노드 탐지 단계이다. 본 논문에서 제안하는 알고리즘은 핫 스팟의 AP에서 동작하면서 이기적 노드 유무를 지속적으로 감시하는 것이 목적이다. 또한 실제 존재하지 않지만 순간적으로 이기적 노드를 검출하는 알고리즘의 오류가 발생할 수도 있다. 이러한 오류를 줄이기 위해서 제안하는 알고리즘의 마지막 단계에서는 이동평균을 이용한 시계열 분석으로 엔트로피 값의 추이를 분석한다.

알고리즘의 3단계에서는 앞서 2단계에서 계산된 뜨살리스-엔트로피 값을 시계열 이동평균 엔트로피 값으로 계산한다. 시계열 이동 평균 모델은 최근 뜨살리스-엔트로피 값을 토대로 직후의 예측 값을 생성하는 방법이다. 식(4)은 데이터 간격 확률변수(BPAT)에 대한 시계열 이동 평균 엔트로피 값으로 계산하는 수식이다.

$$EBPAT_t = \alpha D_t + (1-\alpha) EBPAT_{t-1} \quad (4)$$

여기서, D_t는 현재 t 주기의 $BPAT$ 이동평균 엔트로피, $EBPAT_{t-1}$는 이전 $t-1$ 주기의 $BPAT$ 이동평균 엔트로피, $EBPAT_t$는 현재 $BPAT$ 이동평균 엔트로피 값이다. 그리고 α는 가중치로서 1보다 매우 작은 값을 갖는다. 가중치의 크기는 현재 값을 과거 보다 얼마나 중요시하게 생각하는 가를 나타낸다. 식(5)는 윈도우 구간 최대 데이터 간격(WM) 확률 변수에 대한 이동 평균 엔트로피 값으로 계산하는 수식이다.

$$EWM_t = \alpha D_t + (1-\alpha) EWM_{t-1} \quad (5)$$

여기서, D_t는 현재 t 주기의 WM 이동평균 엔트로피, EWM_{t-1}는 이전 $t-1$ 주기의 WM 이동평균 엔트로피, EWM_t는 현재 WM 이동평균 엔트로피 값이다. 뜨살리스-엔트로피 값은 식(4)와 식(5)를 통해서 각 노드 당 2개의 시계열 이동 평균 뜨살리스-엔트로피 값을 추출하게 된다.

다음으로 두 엔트로피 값을 효과적으로 비교하기 위하여 정규화 하는 작업을 수행한다. 뜨살리스-엔트로피 최댓값은 식 (6)과 같이 정의된다.

$$H_q^{\max} = \frac{1-N^{1-q}}{q-1} \quad (6)$$

여기서, q는 1보다 큰 값이고 N은 확률 변수의 모든 개수이다. 뜨살리스-엔트로피 최댓값을 계산하면 각 파라미터들의 이동 평균 뜨살리스-엔트로피 값을 나누어서 $EBPAT_{norm}, EWM_{norm}$를 계산한다. 정규화는 식 (7), 식 (8)과 같이 정의된다.

$$EBPAT_{norm} = \frac{EBPAT_t}{H(EBPAT)_q^{\max}} \quad (7)$$

$$EWM_{norm} = \frac{EWM_t}{H(EWM)_q^{\max}} \quad (8)$$

여기서, $H_q^{\max}$는 최대 뜨살리스-엔트로피 값이고, EWM_t와 $EBPAT_t$는 각 파라미터의 시계열 이동 평균 뜨살리스-엔트로피 값이다.

제안하는 이기적인 노드 탐지 알고리즘의 최종 단계는 정규화된 뜨살리스-엔트로피 값을 비교하는 작업이다. 알고리즘에서는 각 노드별로 계산 된 두 가지 정규화된 엔트로피 값 중 작은 값을 임계치와 비교하여 이기적 노드 유무를 판단하도록 한다. 이기적인 노드가 이동 작은 엔트로피 값을 가질 것이고 정상 노드는 높은 값을 가질 것으로 예상되기 때문에 더 작은 값을 기준으로 한다. 또한, 임계치는 이기적인 노드를 판별하는데 중요하기에 실제 무선에서 각 환경에 맞도록 임계치를 조절하는 것이 필요하다. 임계치에 따른 실험은 성능 평가에서 이루어진다. 정규화된 엔트로피 값을 사용하여 이기적인 노드를 판별하는 최종 식(9)는 다음과 같이 정의된다.

$$\min[EBPAT_{norm}, EWM_{norm}] < \omega \quad (9)$$

여기서, ω는 이기적 노드 판별을 위한 임계치로 네트워크 환경에 따라 선택할 수 있다. 임계치는 뜨살리스-엔트로피 최댓값보다 작은 값이어야 한다.

5. 성능 평가

본 논문에서는 시뮬레이션을 통해 제안한 뜨살리스-엔트로피를 이용한 이기적인 노드 탐지 알고리즘의 성능 평가를 한다. 시뮬레이터로는 네트워크 시뮬레이터 도구인 ns2(Network Simulator version 2)를 이용한다. 시뮬레이터 상에서 무선 랜 실험환경을 구성하기 위해 표 1의 파라미터를 사용한다.

표 1. 시뮬레이션 파라미터.
Table 1. Simulation parameters.

파라미터	값
Packet Payload	1000 Bytes
MAC Header	28 Bytes
RTS Packet	44 Bytes
CTS Packet	38 Bytes
ACK Packet	38 Bytes
PHY preamble	24 Bytes
Date Rate	11 Bytes
Slot Time	20 us
SIFS	10 us
DIFS	12, 50us
CWmin	8, 16, 32
CWmax	8, 16, 24, 32, 64, 1024

시뮬레이션을 위한 네트워크 모델로는 무선 노드들

과 AP로 단일 무선 랜을 구성하고 무선 노드들이 AP를 통해 유선 노드로 데이터를 전송하는 경우를 가정한다. 다양한 네트워크 혼잡 환경을 고려하기 위하여 무선 랜을 구성하는 무선 노드의 수(n)를 변화시켜 실험을 실시한다. 또한 실험에서는 이기적 노드의 동작방식을 변화시켜 알고리즘의 이기적인 노드 탐지 성공률을 측정한다. 이기적인 노드는 CWmin과 CWmax, DIFS 파라미터를 인위적으로 조정하여 동작하도록 한다. 그림 7는 실험에서 사용된 네트워크 토폴로지를 나타낸 그림이다.

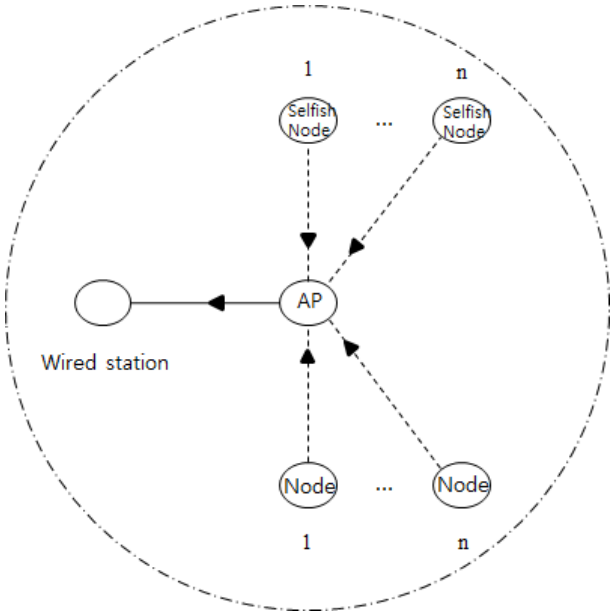

그림 7. 이기적인 노드 탐지 시뮬레이션 토폴로지.
Fig. 7. Simulation topology of selfish node detection.

첫 번째 실험에서는 혼잡도에 따른 이기적 노드 탐지 성공률을 측정하기 위해서 총 노드수를 20개, 16개, 10개, 6개로 변화시켜 시뮬레이션을 실시한다. 그림 8은 총 노드의 개수가 20개인 환경에서의 알고리즘 시뮬레이션 결과이다. 20개의 노드는 이기적인 노드가 7개와 정상 노드 13개를 구성한다. 7개인 이기적인 노드의 동작 방식은 (CWmin 16, CWmax 1024), (CWmin 16, CWmax 32), (CWmin 8, CWmax 1024), (CWmin 8, CWmax 64), (CWmin 8, CWmax 32), (CWmin 8, CWmax 8), DIFS 12us을 각각 적용하여 동작하도록 한다. 첫 번째 시뮬레이션 결과를 통해서 분석할 수 있는 것은 서로 다른 이기적인 노드 7개를 모두 탐지할 수 있는 것이고 이기적인 노드를 동작방식의 등급을 구별할 수 있다는 것이다.

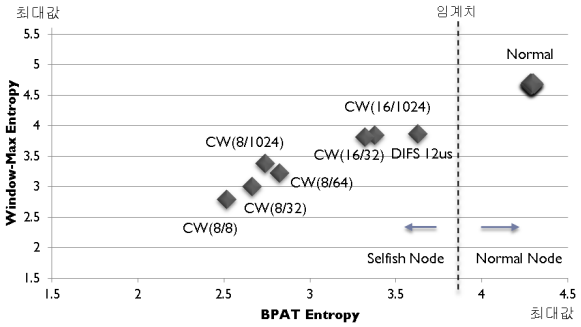

그림 8. 매우 혼잡한 무선 랜 환경(무선 노드 20개)에서 이기적인 노드 탐지 결과.
Fig. 8. Selfish node detection results in case of heavy congested wireless LAN(20 wireless nodes).

두 번째 시뮬레이션에는 중간정도의 혼잡한 환경을 만들기 위해 무선 랜의 총 노드의 수를 16개, 그 중에서 이기적인 노드가 5개, 정상 노드가 11개로 설정한다. 총 5개인 이기적인 노드의 종류로는 (CWmin 16, CWmax 32), (CWmin 8, CWmax 1024), (CWmin 8, CWmax 64), (CWmin 8, CWmax 32), (CWmin 8, CWmax 8)로 각각 다르게 설정한다. 그림 9의 시뮬레이션 결과에서도 역시 이기적인 노드와 정상 노드의 제안하는 뜨살리스-엔트로피 값의 차이가 확연히 구분됨을 확인할 수 있으며, 정상 노드의 경우에 엔트로피 유사한 값을 얻을 수 있는 것을 볼 수 있다.

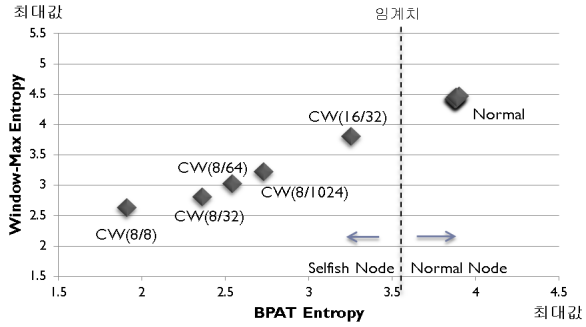

그림 9. 중간정도 혼잡한 무선 랜 환경(무선 노드 16개)에서 이기적인 노드 탐지 결과.
Fig. 9. Selfish node detection results in case of normally congested wireless LAN(16 wireless nodes).

세 번째 시뮬레이션은 무선 랜에서 약한 혼잡을 만들기 위해 무선 랜에서 총 노드의 수를 10개, 그 중에서 이기적인 노드가 4개, 정상 노드가 6개로 설정한다. 총 4개인 이기적인 노드의 종류로의 파라미터는 각각 (CWmin 8, CWmax 1024), (CWmin 8, CWmax 64), (CWmin 8, CWmax 32), (CWmin 8, CWmax 8)로 설정한다. 그림 10의 시뮬레이션 결과는 앞의 두 경우와 유사한 결과를 얻을 수 있다. 즉, 이기적인 노드와 정상 노드의 뜨살리스-엔트로피 값의 차이가 임계치를 중심으로 확연히 구분됨을 확인할 수 있다.

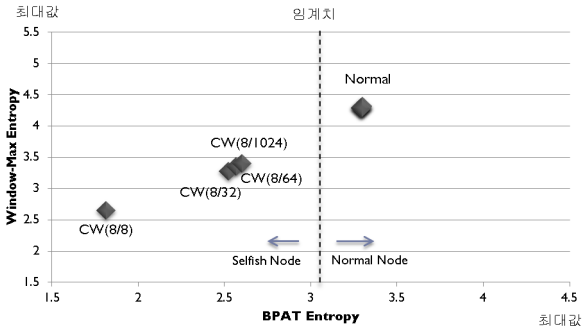

그림 10. 약하게 혼잡한 무선 랜 환경(무선 노드 10개)에서 이기적인 노드 탐지 결과.
Fig. 10. Selfish node detection results in case of lightly congested wireless LAN(10 wireless nodes).

네 번째 시뮬레이션에 환경은 무선 랜이 혼잡하지 않는 상황에서 이기적인 노드를 탐지에 대한 실험을 수행한다. 이를 위해 총 노드의 수를 6개, 그 중에서 이기적인 노드 3개, 정상 노드 3개로 실험환경을 구성한다. 사실 혼잡하지 않는 경우 이기적 노드에 의해 정상 노드의 전송률이 저하되지 않지만 그림 11의 실험결과에서 확인할 수 있듯이 제안하는 알고리즘은 성공적으로 이기적인 노드 3개를 100%로 탐지할 수 있다.

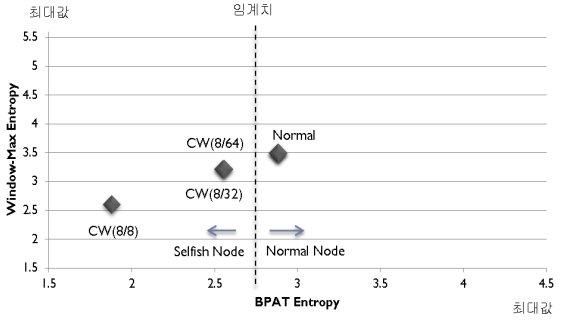

그림 11. 혼잡하지 않은 무선 랜 환경(무선 노드 6개)에서 이기적인 노드 탐지 결과.
Fig. 11. Selfish node detection results in case of uncongested wireless LAN(6 wireless nodes).

무선 네트워크에서의 환경은 언제나 일정할 수 없으며 무선 노드의 수가 균등하게 유지되는 환경이 아니기 때문에 정상 노드의 수와 이기적인 노드의 수가 다양할 수 있다. 그래서 무선 환경에서 무선 노드의 수를 다양하게 실험을 해야 할 필요가 있고 이기적인 노드의 수 역시 점점 늘려가면서 탐지 성공률을 확인해야 한다.

따라서 두 번째 실험은 무선 노드를 2개부터 20개까지 늘려가면서 실험을 실시하고, 논문에서 제안하는 엔트로피 값의 변화를 관찰한다. 이 실험에서 확인할 수 있는 것은 각 파라미터 별로 이기적인 노드와 정상 노드를 구분하는 엔트로피 값의 변화이다. 특히 Window-Max(WM) 파라미터의 경우에는 혼잡하지 않는 상황과 매우 혼잡한 상황의 두 가지 경우에서 BPAT 파라미터보다 더 낮은 탐지 성공률을 보여주는 것을 확인 할 수 있다 (그림 12).

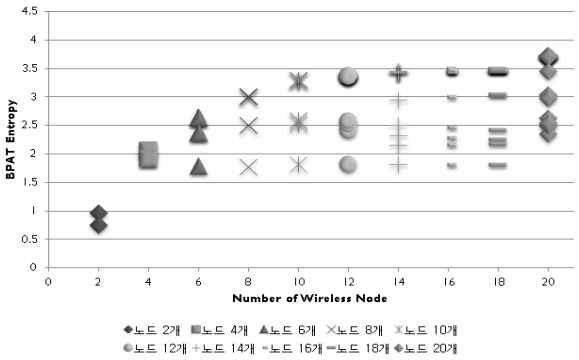

그림 12. 노드 수 변화에 따른 BPAT-Entropy 값.
Fig. 12. BPAT-entropy value according to varied number of wireless nodes.

이것을 통해서 기존의 BPAT 파라미터만을 수집하지 않고 Window-Max 파라미터를 추가적으로 수집하여 엔트로피로 변환한 목적이 실험을 통해서 증명되었다. 결과적으로 실험은 정상 노드들은 수치상으로 밀집되어 있었고 제안하는 알고리즘을 통해서 이기적인 노드들이 정상 노드와 구분이 가능하다는 것을 확인하였다. 그림 13은 각 환경마다 Window-Max Entropy의 변화를 측정한 그래프이다.

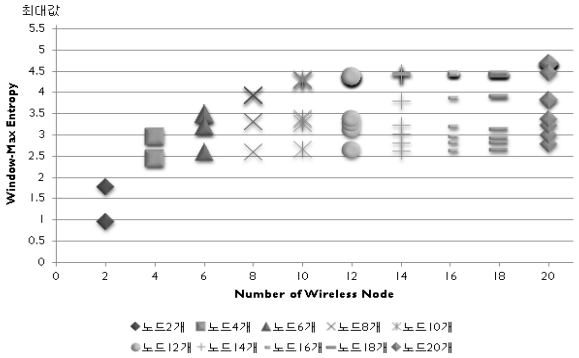

그림 13. 노드 수 변화에 따른 Window-Max Entropy.
Fig. 13. Window-Max entropy value according to varied number of wireless nodes.

제안하는 알고리즘에서 이기적인 노드와 정상 노드를 판단할 때 중요하게 사용되는 것은 엔트로피 임계치이다. 이 임계치는 무선 환경의 상황에 맞게 달라질 수 있고 본 실험에서 각각 다른 임계치를 사용하여 이기적인 노드 탐지 성공률을 측정해보았다. 탐지 성공률에서 100%는 미리 설정한 이기적인 노드를 모두 탐지한 것을 의미하며, 85%, 80%, 57% 등은 총 이기적인 노드 중 탐지하지 못한 노드를 제외한 것을 의미한다.

첫 번째 실험은 무선 노드 20개에서 이기적인 노드를 7개로 설정하여 각 노드가 AP에게 전송을 하도록 실험하였다. 그렇게 얻은 실험 결과는 뜨살리스-엔트로피로 변환해 본 결과의 수치를 임계치와 비교하였고 이기적인 노드로 판단할 때 각각 다른 임계치에 따라 이

기적인 노드 탐지 성공률을 측정해보았다. 실험을 통해서 이기적인 노드가 존재하는 무선 환경에서 정상 노드들의 뜨살리스-엔트로피는 밀집되어 있으며 큰 차이가 나지 않는 특징을 보여주었고 엔트로피 수치가 높게 측정되었다. 임계치를 뜨살리스-엔트로피 최댓값의 50%, 60%, 70%, 80%로 변환해서 탐지 성공률을 측정해본 결과가 표 2와 같다. 노드 20개 중 이기적인 노드가 7개인 경우에서 임계치 70%와 80%가 이기적인 노드를 모두 탐지하였고 50%와 60%인 경우에 탐지하지 못하는 결과가 나타났다. 두 번째 실험은 무선 노드 15개 중에서 이기적인 노드를 5개로 설정하였고 서로 다른 임계치에 따라서 이기적인 노드 탐지 성공률을 측정하였다. 표 3은 두 번째 실험 결과이고 여기에서 임계치가 3.36(60%) 이상에서는 탐지 성공률이 성공적으로 100%가 나왔으며 2.8(50%)에서는 80%의 탐지 성공률을 보여주었다. 두 번째 실험에서도 첫 번째 실험과 같이 70%이상의 임계치의 성공률이 동일하게 나왔다. 세 번째 실험은 무선 노드 10개 중에서 이기적인 노드를 4개로 설정하였고 일반적인 무선 네트워크에서의 환경을 모델로 실험하였다. 표 4에서 세 번째 실험 결과이고 임계치가 60%에서 80%까지 100%의 성공률을 보여주었고 2.8(50%)에서 오탐률을 보여주었다. 네 번째 실험은 무선 환경에서 극한 환경을 모델로 하였으며 혼잡하지 않은 환경을 구성하기 위해서 총 무선 노드의 수를 5개로 지정하였고 이기적인 노드는 2개로 설정하였다. 이러한 설정으로 실험을 하였을 때 노드들이 무선 채널을 전송량 기준에서 볼 때 전부 사용하지 않았고 통계적인 기법을 사용하여 이기적인 노드와 정상 노드를 구분하기 쉽지 않다. 하지만 표 5는 네 번째 실험 결과이고 이 표에서 볼 수 있듯이 BPAT와 Window-MAX 파라미터에서 탐지 성공률이 각각 다른 임계치에서 100%의 탐지 성공률을 보여주었다. 실험 분석 결과 제안하는 알고리즘은 이기적인 노드를 탐지하는 탐지 성공률에 대해서 좋은 성능을 보여준 것으로 판단된다.

표 2. 매우 혼잡한 무선 랜 환경(무선 노드 20개)에서 임계치별 이기적인 노드 탐지 성공률
Table 2. Selfish node detection success rate per threshold value in case of heavy congested wireless LAN(20 wireless nodes).

임계치 ω	탐지 성공률(%)
4.48 (80%)	100 %
3.92 (70%)	100 %
3.36 (60%)	85 %
2.8 (50%)	57 %

표 3. 중간정도 혼잡한 무선 랜 환경(무선 노드 15개)에서 임계치별 이기적인 노드 탐지 성공률
Table 3. Selfish node detection success rate per threshold value in case of normally congested wireless LAN(15 wireless nodes).

임계치 ω	탐지 성공률(%)
4.48 (80%)	100 %
3.92 (70%)	100 %
3.36 (60%)	100 %
2.8 (50%)	80 %

표 4. 약하게 혼잡한 무선 랜 환경(무선 노드 10개)에서 임계치별 이기적인 노드 탐지 성공률
Table 4. Selfish node detection success rate per threshold value in case of lightly congested wireless LAN(10 wireless nodes).

임계치 ω	탐지 성공률(%)
4.48 (80%)	100 %
3.92 (70%)	100 %
3.36 (60%)	100 %
2.8 (50%)	80 %

표 5. 혼잡하지 않은 무선 랜 환경(무선 노드 5개)에서 임계치별 이기적인 노드 탐지 성공률
Table 5. Selfish node detection success rate per threshold value in case of uncongested wireless LAN(5 wireless nodes).

임계치 ω	탐지 성공률(%)
4.48 (80%)	100 %
3.92 (70%)	100 %
3.36 (60%)	100 %
2.8 (50%)	100 %

6. 결 론

최근 급격히 증가하는 무선 랜 핫 스팟에서 인위적인 방법으로 정상 노드보다 전송 데이터 량을 늘리려고 하는 이기적 노드 문제가 최근 연구되고 있다. 본 논문에서는 기존의 이기적인 노드 탐지를 위한 프로토콜 기반의 방법에서 탈피하여 통계적 기법인 뜨살리스-엔트로피를 이용한 이기적인 노드 탐지 알고리즘을 제안하였다. 제안하는 알고리즘에서는 AP에서 주기마다 반복적으로 동작하며, 노드별 데이터 간격과 윈도우 구간 최대 데이터 간격을 확률 변수로 하는 뜨살리스-엔트로피를 값을 도출한 후 이동평균, 정규화, 그리고 임계치와의 비교를 통해 이기적 노드 유무를 판별한다. 마지막으로 논문에서는 제안한 이기적인 노드 검출 알고리즘의 성능을 평가하기 위하여 다양한 무선 랜 환경(혼잡도, 이기적 노드 동작방법, 임계치)등을 고려하여 시뮬레이션을 수행하였고, 그 결과 제안한 방법의 이기적 노드 탐지 성공률이 매우 우수 하다는 것을 확인하였다.

참 고 문 헌

[1] IEEE 802.11 Working Group, "IEEE 802.11-2007: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," June 2007.

[2] M. Raya, J. P. Hubaux, and I. Aad, "DOMINO: Detecting MAC layer greedy behavior in IEEE 802.11 hotspot," *IEEE Transaction on Mobile Computing*, vol. 5, issue. 12, pp. 1681-1705, Dec. 2006.

[3] Konstantinos Pelechrinis, Guanhua Yan., etc. "Detecting Selfish Exploitation of Carrier Sensing in 802.11 Networks," *IEEE Infocom 2009 proceedings*, pp. 657-665, 2009.

[4] Pablo Serrano, Albert Banchs, Valerio Targon, and Jose Felix Kukjelka, "Detecting Selfish Configurations in 802.11 WLANs," *IEEE Communications Letters*, vol. 14, issue. 2, pp. 142-144, 2010.

[5] 한유훈, 강동훈, 석승준, "통계적 분석을 통한 무선 랜의 탐욕(Greedy) 노드 탐지," *통신망 운용 관리 학술 대회*, 2011.

[6] M. Cagalj, S. Ganeriwal, I Aad, and J.P Hubaux, "On Selfish Behavior in CSMA/CA Networks," *in Proc. of the IEEE INFOCOM*, vol. 4, pp. 2513-2524, 2005.

[7] C.E. Shannon, "Prediction and entropy of printed English," *The Bell System Technical Journal*, vol. 30, pp. 50-64, 1951.

[8] Giseop No, Ilkyeun Ra, "An Efficient and Reliabl e DDoS Attack Detection Using a Fast Entropy Computation Method," *Communications and Information Technology*, pp. 1223-1228, Sept. 2009.

[9] Qian Quan, Che Hong-Yi, Zhang Rui, "Entropy Based Method for Network Anomaly Detection," *IEEE Pacific Rim International Symposium on Dependable Computing*, pp. 189-191, 2009.

[10] Artur Ziviani, Antonio Tadeu A. Gomes, Marcelo L. Monsores, Paulo S. S. Rodrigues, "Network Anomaly detection using Nonextensive Entropy," *IEEE Communication Letters*, vol. 11, no. 12, pp. 1034-1036, 2007.

저 자 소 개

류 병 현 (Byoung-Hyun Ryu)

2010년 : 경남대학교 컴퓨터공학부 졸업(공학사).

2010년 3월 ~ 현재 : 경남대학교 대학원 첨단공학과 재학(공학석사)

관심분야 : 컴퓨터 네트워크, 네트워크 보안

Phone : 070-7527-3590

Fax : 055-248-2554

E-mail : temple2@naver.com

석 승 준 (Seung-Joon Seok)

2003년 : 고려대학교 대학원 전자공학과 졸업(공학박사).

2004년 ~ 현재 : 경남대학교 컴퓨터공학부 부교수

관심분야 : 무선인터넷 프로토콜, 미래인터넷, USN 프로토콜 etc.

Phone : 055-249-2710

Fax : 055-248-2554

E-mail : sjseok@kyungnam.ac.kr