센서네트워크에서 유전자 알고리즘을 이용한 침입탐지시스템 노드 스케줄링 연구

성기택*

A Study on the Intrusion Detection System's Nodes Scheduling Using Genetic Algorithm in Sensor Networks

Ki-taek Seong*

요 약

센서네트워크의 다양한 응용분야에서 보안성은 대단히 중요하다. 침입탐지는 공격에 대한 방어기법 중의 하나이지만 기존의 정형화된 침입탐지기술은 제한된 자원으로 운영되는 센서네트워크에는 적절하지 않다. 본 논문에서는 전송되는 패킷의 이상행위를 관찰하는 침입탐지시스템에서 탐지노드의 선정 및 운영에 관한 방법과 함께 침입탐지시스템의 수명을 최대화하는 노드 스케줄링 방안을 제안하였다. 제안된 최적화식에 대하여 유전자 알고리즘을 이용한 해를 개발하고 시뮬레이션을 수행하여 효율성을 확인하였다.

ABSTRACT

Security is a significant concern for many sensor network applications. Intrusion detection is one method of defending against attacks. However, standard intrusion detection techniques are not suitable for sensor networks with limited resources. In this paper, propose a new method for selecting and managing the detect nodes in IDS(intrusion detection system) for anomaly detection in sensor networks and the node scheduling technique for maximizing the IDS's lifetime. Using the genetic algorithm, developed the solutions for suggested optimization equation and verify the effectiveness of proposed methods by simulations.

키워드

센서네트워크, 침입탐지, 보안, 최적화, 유전자 알고리즘

Key word

sensor networks, intrusion detection, security, optimization, genetic algorithm

* 정회원 : 동명대학교(ktseong@tu.ac.kr)

접수일자 : 2011. 08. 17
심사완료일자 : 2011. 09. 20

Ⅰ. 서 론

1.1. 센서네트워크와 보안

무선통신 및 전자관련 기술의 발달에 따라 저가의 저 전력 다중센서노드 구현이 가능하게 되었다. 감지, 데이터 처리, 그리고 통신모듈들의 소형화에 따라 센서네트워크를 구성하는 노드의 소형화도 가능하게 되었다. 센서네트워크는 전장에서 전력화 요소들의 이동을 탐지, 환경오염의 감시, 도로상에서의 교통량 모니터링, 건물에서의 사람들의 위치추적 등 다양한 분야에서의 이용되고 있다. 이와 같은 응용분야의 특수성으로 인하여 센서네트워크에서는 높은 보안성이 요구된다[1-2]. 센서네트워크의 특성에 따른 보안요소는 다음과 같다[3].

- 데이터 비밀성/무결성/신선성
- 자율구성(self-organization)
  - 다중 hop 라우팅을 지원
  - 보안 채널형성과 유지를 위한 키분배 지원
- 시간 동기성(time synchronization)
- 보안의 국지성과 보안

센서네트워크는 환경에 노출되어 분포되므로 물리적 공격에 무방비 상태이며 이로 인하여 공격받기 쉽다. 또한 제한된 에너지로 인하여 노드의 전력을 소비시키거나 무선 통신을 방해하는 형태의 공격이 가능하다. 가장 주목할 공격형태는 DOS(denial-of-service)로서 이에 대하여 네트워크 계층별 대응기술이 연구되어 왔다. 그 이외에 암호화를 통한 보안접근방법으로서 공개키 암호화방식, 대칭키 암호화방식이 센서네트워크에 적용되었으며, 이를 위한 키 관리 프로토콜로서는 중앙관리/분산관리, 키 분배 확률 기반의 프로토콜 등의 기술이 적용이 이루어져왔다. 또한 보안 라우팅 기술을 적용한 방식과 보안 데이터 수집기술 등이 제안되었다. 이상과 같은 방식은 외부로부터의 공격을 방어하기 위한 것이며 이미 외부로부터 침입을 당하였을 경우 이를 탐지하는 기능도 중요하여 이를 지원하는 침입탐지시스템에 대한 연구도 진행되고 있다[6].

본 연구에서는 별도의 침입탐지용 부가장치를 고려하지 않고 기 분포된 노드를 이용한 침입탐지장치구축방법에서 에너지 효율적인 운영기법을 제안하였다. 논문의 구성은 먼저 센서네트워크에서의 IDS 소개와 관련연구를 소개한다. 3장에서는 에너지 효율적인 IDS 운영기법의 모델링과 이에 대한 알고리즘에 대하여 기술하고 4장에서는 효율성 검정을 위한 시뮬레이션 결과를 함께 5장에서 연구결론을 기술한다.

Ⅱ. 센서네트워크 침입탐지

2.1. 센서네트워크에서 IDS

전술한바와 같이 센서네트워크에서 외부로부터의 침입을 방어하는 것도 중요하지만 이미 침입을 허용했을 때 이를 탐지하는 기술도 중요한다. 센서네트워크용 IDS는 이상행위(anomaly)의 탐지(AID : anomaly based intrusion detection)와 자원의 악용(misuse)을 탐지(MID : misuse intrusion detection)하는 것으로 분류된다[7]. 센서네트워크 IDS의 기능적 구조를 그림 1에 나타내었다.

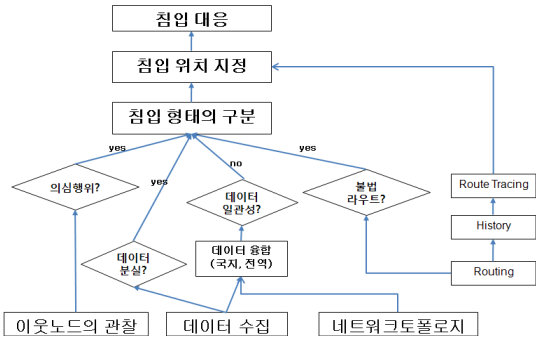

그림 1. 센서네트워크 IDS
Fig. 1 IDS in Sensor Networks

그림 1에 나타난 바와 같은 IDS프레임워크는 다양하게 구성될 수 있으며 접근하는 방법에 따라 다음과 같이 기능을 구현한다.

- 이웃하는 노드가 특정 노드를 관찰
- 네트워크에서 전송되는 데이터를 분석
- 네트워크를 구성할 때 라우팅 정보의 변화에 따른 토폴로지 변화를 관찰

각각의 접근하는 방법에 따라 다양한 방법들이 연구되고 있으며, 본 연구는 이웃노드의 관찰을 통한 IDS구

현 및 운영에 관한 것이다.

2.2. 관련 연구

센서네트워크 IDS의 구현기술에 관한 연구가 최근까지 많이 이루어져왔다. Anjum 등은 최소 컷 집합(minimum cut-set)과 최소 영역 집합(minimum dominating-set) 개념을 이용하여, 특별한 탐지 하드웨어를 장착한 노드의 수를 최소화하여 분포시키는 방법을 제안하였으며[10], Agah 등은 네트워크 트래픽 부하와 Markov 결정 프로세스를 기반으로 하는 게임기법이 탐지기법에 적용할 수 있음을 증명하였으며[11], Loo 등은 클러스터링 알고리즘과 이상행위 탐지기법을 이용한 IDS 구현방법을 제안하였고[12], Su 등은 클러스터 기반의 네트워크에서 이웃 노드로 전송되는 비정상적인 패킷을 인식토록 함과 동시에 클러스터헤더를 모니터하여 침입탐지기법에 적용하는 방안을 제시하였다[13]. Roman 등은[9]에서 "spontaneous watchdog" 개념을 사용하여 이웃하는 노드를 패킷 탐지노드로 설정하는데 하나의 패킷에 대하여 최소 하나의 탐지노드가 설정되도록 확률적인 모델을 제안하여 IDS에 적용하였다. [15]에서는 IDS 노드의 공격에 대한 감지확률과 에너지 효율성에 대한 연구를 나타내었다. 센서네트워크에서의 IDS에 관하여 기존의 연구에서 이웃하는 노드를 이용한 경우 에너지 효율을 고려한 운영기법에 관한 연구는 적다.

Ⅲ. 제안하는 방법

본 연구에서 제안하는 IDS 구성과 운영방법을 요약하면 다음과 같다. 먼저 센서네트워크가 BS(base station) 기반으로 운영된다고 가정할 때 BS에 의하여 라우팅이 완료되면 센서네트워크가 형성된다. 이때 네트워크의 패킷을 관찰하는 IDS는 네트워크에 참여하지 않은 유효노드를 이용하여 구성되도록 한다. 전체 네트워크를 커버하는 IDS를 구성하는 노드들을 선택하는데 있어서 에너지 효율적인 방법을 다음과 같이 제안한다.

3.1 탐지노드

3.1.1 탐지노드기능

탐지노드의 기능으로서 이웃하는 노드들의 송수신 패킷의 분석을 통한 네트워크 자체 외부침입여부를 판단하여 외부로 알려주는 시스템을 고려하였다. 그러므로 별도탐지 전용노드를 사용하지 않고 수신되는 패킷의 주소부분만 검토하여 특정 노드로 패킷의 편중현상이 발생하는지의 여부만 인식하는 기능은 소프트웨어만으로 충분히 가능하므로 에너지 소모에 큰 영향을 미치진 않는다. 센서노드는 일반적으로 에너지 효율적인 운영을 위하여 활동상태(active mode)와 유휴상태(idle mode)를 구분하여 적절한 동작모드로 운영된다. 그리고 노드가 유휴상태에 있더라도 무선 수신은 가능하다[14].

본 연구에서는 수신/유휴상태에 있는 노드를 탐지노드로 사용하는 방법을 고려하였다. 탐지노드로서 에너지소모에 있어서 일반노드와 비교시 부가되는 에너지는 수신된 패킷의 분석에 필요한 CPU 부하에 불가하여 미비하다. 예를 들면 sinkhole 공격의 경우 패킷의 전송방향만으로도 침입여부를 판단할 수 있다.

3.1.2 탐지노드 선정방법

센서네트워크에서는 일반적으로 주어진 임무에, 예를 들어 특정지역에서의 화재감지, 필요한 노드 이상의 수로 분포되어 운영된다. 따라서 잉여의 노드가 발생하게 되며 이러한 잉여노드는 고장대체, 에너지 효율성 제고를 위한 임무교대, 고장극복(fault recovery) 등의 용도로 사용된다.

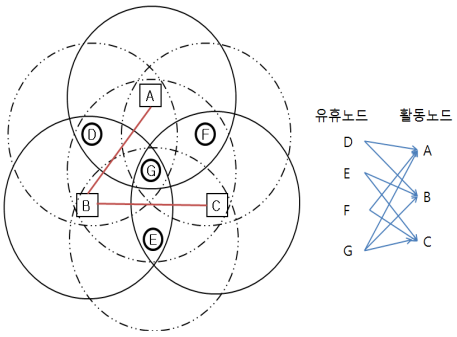

그림 2. 탐지노드 선정 원리
Fig. 2 selection principle for detector node

그림2는 7개의 노드가 분포된 상태에서 3 개의 노드(Ⓐ, Ⓑ, Ⓒ)는 네트워크의 일부로 활용되어 패킷이 전송되며 나머지 노드(Ⓓ, Ⓔ, Ⓕ, Ⓖ)는 잉여노드이다. 그림 2

에 의하면, 노드 A에서 노드 B를 경유하여 C로 가는 경로상의 모든 패킷을 수신하는 탐지노드가 되기 위한 조건은 경로를 커버하는 통신반경을 갖는 위치에 있는 노드이다. 이는 곧 A, B, C노드의 통신영역에 존재하는 모든 잉여노드는 감지기 역할이 가능함을 의미한다.

3.1.3 탐지노드 선정문제

모델링을 위하여 다음과 같이 가정한다.

1) 이미 충분한 수의 노드가 랜덤하게 분포되어있고 그 중의 일부가 현재 활동 중인 노드들로 구성된 센서네트워크가 있으며 나머지 노드는 유휴노드로 대기하고 있다.
2) 센서네트워크는 BS기반으로 운영되며 이웃하는 다른 센서노드로부터의 데이터 수집 기능을 수행한다.
3) BS는 산하 노드의 에너지준위, 위치에 대한 정보를 갖고 있으며 라우팅과 함께 노드의 동작모드를 결정하는 등의 네트워크 운영을 담당한다.

그림 2에서와 같이 경로를 탐지하는 노드의 집합은 하나의 IDS 역할을 할 수 있으며 그림 2에서, 각 유휴노드는 다음과 같이 네트워크 노드들을 커버한다. 즉, Ⓓ ={Ⓐ, Ⓑ}, Ⓔ={Ⓑ, Ⓒ}, Ⓕ ={Ⓐ, Ⓒ}, Ⓖ ={Ⓐ, Ⓑ, Ⓒ}

3.2. 에너지 효율적인 운영모델

센서노드로 사용되는 MICA2의 경우 유휴/수신 상태에서 에너지 소비는 동일하다[14]. 따라서 IDS의 패킷수신상태에서의 에너지 소비량과 일반 유휴상태에서의 에너지소비에서의 차이는 미미하다. 그러나 일단 이상이 탐지되면 IDS는 이 사실을 알려야하므로 IDS를 구성하는 탐지노드는 활성화되어 알림패킷을 BS까지 전송해야한다. 따라서 본 연구에서는 IDS가 이상을 탐지하여 알림패킷을 전송하기까지의 과정에서 소모되는 에너지소비 형태를 고려하였다.

모든 유휴노드의 잔량 에너지는 단위시간 '1'이라 가정하면 전체 노드가 연속적으로 동작할 때 네트워크의 수명은 '1'이 된다.

[16]에 의하면 그림 2와 동일한 조건에서 교집합을 허락하지 않는 조건(disjoint)에서 전체를 커버하는 집합은 $S_1 = \{D, E\}$, $S_2 = \{F, G\}$이며, 결과적으로 IDS의 수명은 '2'가 된다. 그러나 이러한 조건이 아닐 경우에는 IDS 수명이 개선될 수 있다. 예를 들면 S_1 ={Ⓓ, Ⓔ}에 '0.5', S_2 ={Ⓔ, Ⓕ}에 '0.5', S_3={Ⓓ, Ⓕ}에 '0.5' 그리고 S_4={Ⓖ}에 '1'을 적용하면 IDS의 수명은 '2.5'가 된다.

이와 같이 탐지노드에 적절한 시간으로 운영할 경우 IDS의 수명은 개선될 수 있다. 이와 같은 사항을 고려하여 전체 네트워크를 커버하는 IDS를 구성하는 노드의 집합을 구하는 문제를 정의하면 다음과 같다.

IDS 집합문제 : 유휴 센서노드 집합 C와 현재 활동 중인 노드 집합 R이 주어졌을 때 전체 R을 커버하는 센서들의 집합 $S_1, S_2, \cdots, S_p$와 이에 대응하는 $t_1, t_2, \cdots, t_p$를 구하되 $t_1 + t_2, \cdots, + t_p$의 값이 최대가 되도록 한다.

IDS집합문제는 최적화문제이며 정수 프로그래밍(integer programming) 형으로 모델링하면 다음과 같다.

$$\text{Maximize} \quad t_1 + \cdots + t_p \tag{1}$$

$$\text{subject to} \sum_{j=1}^{p} x_{ij} t_j \le 1 \; \forall \, s_i \in C$$

$$\sum_{i \in C_k} x_{ij} \ge 1 \; \forall \, r_k \in R, \, j = 1, \cdots, p$$

$$\text{where} \; x_{ij} = 0,1 \; (x_{ij} = 1 \; \text{if and only if} \; s_i \in S_j)$$

여기서,p를 활동 네트워크를 커버하는 집합의 수이고 n개의 유휴노드 집합 $C = \{s_1, s_2, ..., s_n\}$, m개로 구성된 활동 중인 네트워크 노드 집합 $R = \{r_1, r_2, ..., r_m\}$, 이라할 때 유휴노드와 활동노드와의 관계 $C_k = \{i \mid sensor \; s_i \; covers \; active \; node \; r_k\}$ 로 되며, 변수 x_{ij}는 이진 값 (i=1..n, j=1..p;)으로 만일 s_i가 S_j를 구성하는 노드이면 1, 그렇지 않으면 0의 값을 가지며, $t_j \in \mathcal{R}, 0 \le t_j \le 1$으로 집합$S_j$에 할당되는 시간을 의미한다.

3.3. IDS 집합문제의 해

3.2에서 정의된 IDS집합문제는 최적화식으로서 다양한 해결방안이 고려될 수 있다.

3.3.1 Greedy 기법

Greedy 기법은 항상 최적의 해를 보장하지 않으나 간단하며 빠른 수행속도에서의 장점을 갖고 있다. 일반적으로 센서네트워크는 대량의 노드가 사용되므로 노드의 수에 따라 최적화 방정식의 계산 량은 크질 수 있으므로 연산 량의 부하를 감소한다는 장점으로 Greedy 알고리즘을 사용하였으며 해의 순서는 다음과 같다.

```
Greedy (C, R, w)
1: set lifetime of each sensor to 1
2: SENSORS = C
3: i=0
4: while each target is covered by at least one sensor
   in SENSORS do
5:   /*a new set cover Ci will be formed */
6:   i = i + 1
7:   Ci = ∅
8:   TARGETS = R
9:   while TARGETS = ∅do
10:    /*more targets have to be covered */
11:    find a critical target r_critical ∈ TARGETS
12:    select a sensor s_u ∈ SENSORS with greatest
       contribution, that covers rcritical
13:    Ci = Ci ∪ s_u
14:    for all targets r_k ∈ TARGETS do
15:      if r_k is covered by s_u then
16:        TARGETS = TARGETS −r_k
17:      end if
18:    end for
19:  end while
20:  for all sensors s_j ∈ Ci do
21:    lifetime s_j = lifetime s_j −w
22:    if lifetime s_j == 0 then
23:      SENSORS = SENSORS - s_j
24:    end if
25:  end for
26: end while
27: return i-number of set covers and the set covers C1,
C2, ..., Ci
```

알고리즘에서 C는 모든 노드집합, R은 감시해야할 활동 중인 노드의 집합, w는 노드의 수명 값이며 한번 관찰노드로 활동하면 $w \in (0,1]$의 크기만큼 감소된다. TARGETS은 관찰해야할 활동 중인 노드 집합이다. 모든 활동노드 커버여부는 줄 14에 나타낸 바와 같이 기본적으로 모든 노드를 커버하는 집합만을 대상으로 하기 때문에 전역 커버를 기본적으로 만족한다. 줄 5부터 19까지가 발견적 기법이 적용되었다. 11의 $r_{critical}$은 이웃하는 유휴노드들의 개수와 잔존에너지 고려하여 선택하였다. 왜냐하면 이렇게 함으로서 유휴노드를 균등하게 활용하여 전체 에너지 소비를 효율적으로 하는 효과를 기대할 수 있기 때문이다. 하나의 C_i가 구해지면 C_i를 구성하는 모든 노드의 에너지는 w만큼 감소되며 다시 처음으로 돌아가 다음의 C_i를 구하게 된다.

3.3.2 IDS 노드 동작 스케줄링

Greedy 기법에 의하면 하나의 C_i가 구해지면(알고리즘 줄9 - 19) 이에 따라 IDS구축되어 임무에 들어가게 되며 일정 에너지를 소모하고(알고리즘 줄20 - 26) 다음 C_i를 구하기 위하여 되돌아간다. 즉, IDS의 관찰노드는 스케줄에 따라 유휴/활동상태 전환이 발생하게 되는데 이때 C_i를 구하는 시간동안 탐지노드는 유휴상태로 되어 이상 현상이 발생되더라고 전달할 수 없게 된다.

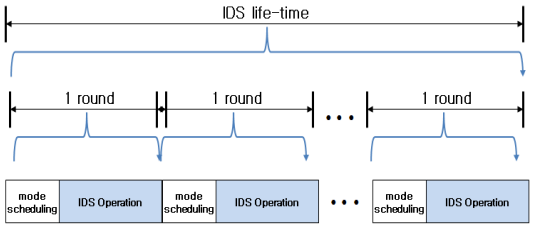

그림 3. Greedy 알고리즘에서의 스케줄링
Fig. 3 scheduling for Greedy algorithm

그림 3에서스케줄링 시간은 하나의 집합 C_i를 구하는 시간이며 동작시간은 C_i에 따라 IDS로 동작하는 시간이다. 이 두 가지 동작을 라운드 시간이라 한다면 라운드에 존재하는 스케줄링 시간동안 이상 현상이 발생하더라도 이를 즉시에 전달할 수 없게 되는 단점이 있다. 이는 IDS의 지속적인 관찰이라는 QoS (Quality of Service)관점에서 중요한 요인이라 할 수 있다.

3.4. 유전자 알고리즘 기법

유전자 알고리즘 기법도 최적화문제 해결방안으로 사용될 수 있다. 본 연구에서는 지속적인 IDS기능을 보장하기 위한 방법으로 유전자 알고리즘 기법을 제안한다. 이를 위하여 전술한 IDS구성요소 집합 C_i를 한 번에 구하여 이를 순차적으로 적용함으로서 연속적인 임무수행을 가능하게 하는 장점을 제공한다. 유전자 알고리즘의 일반적인 동작순서는 그림 4와 같다. 그림에서 종료조건(termination condition)을 만족하는 모든 집합 C_i 집합을 단일 수행을 통하여 구할 수 있으므로 이를 적용하면 그림 5와 같은 IDS 노드 스케줄링이 가능하다. 즉 연속적인 관찰이 가능하게 되므로 관찰의 연속성이란 QoS 이점이 있다.

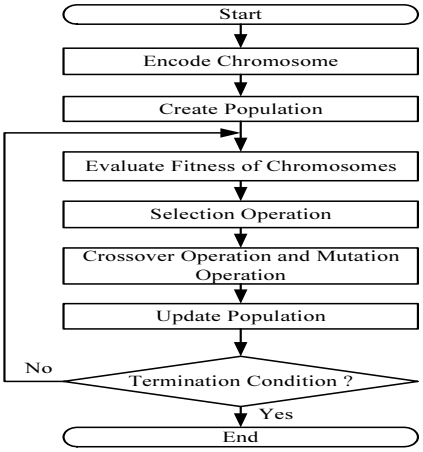

그림 4. 유전자 알고리즘의 동작
Fig. 4. The operation of genetic algorithm

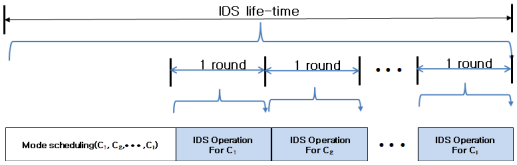

그림 5. 유전 알고리즘에서의 스케줄링
Fig. 5 scheduling for Genetic algorithm

본 연구에서는 중앙처리방식의 유전자 알고리즘 기법을 이용하여 IDS집합문제를 해결하고자 하였다. 즉 BS가 모든 노드에 대한 위치, 에너지 준위 값을 알고 있으므로 이를 이용한다. 유전자 알고리즘을 이용하여 IDS집합문제를 해결하기 위하여 문제에 적합한 염색체 표현방법과 함께 새로운 유전연산자들을 설계한다. 주어진 센서노드와 목표지점들에 대한 정보를 알고 있을 때, 염기는 모든 목적지점을 커버하는 하나의 노드들의 집합을 의미하며, 염색체는 IDS집합문제의 잠재 해로서 염기들의 집합을 의미한다.

3.4.1 염색체의 표현

유전자 알고리즘에서 문제의 잠재 해를 표현하기 위한 기본적인 개체를 염색체라 하고 각 개체를 이루는 요소를 염기라 한다. 유전자 알고리즘에서 모든 해는 염색체로 표현될 수 있어야 하므로, 가장 보편적으로 사용하는 염색체의 표현(또는 인코딩)방식은 이진비트열을 이용한 방법이다. 이진비트열 방식은 가능한 모든 해를 표현할 수 있어야 하므로 주어진 문제에 대한 모든 해의 구성요소가 결정된 상태에서 적용가능하다. 커버해야할 활동노드를 "target"이라 지칭하여, 본 연구에서 제안한 방법에서는, 염색체를 부호화하기 위해서 미리 모든 target 커버를 수행하는 센서들의 집합들을 구하지 않고, 센서들을 하나씩 임의로 선택하여 모든 target이 커버될 때 까지 반복 수행함으로써 하나의 염기를 구성하게 하고, 이러한 염기들로 구성된 염색체를 문제의 해로 표현한다. 염색체의 인코딩 과정을 그림5에 나타내었다. 그림에서 $S_1, S_2, S_3, \cdots S_p$ 는 모든 target을 커버하는 센서노드 집합인 염기에 해당된다. 여기서 S_i는 앞서 언급한 바와 같이 미리 구하지 않으며, 임의의 센서노드를 선택하여 target 커버여부를 결정하게 되므로, S_i 집합의 크기, 구성요소 그리고 집합의 수는 다르게 나타날 수 있다. 실제 $S_1, S_2, S_3, \cdots S_p$ 는 염색체 인코딩 과정에서 만들어지며 이러한 절차에 따라 각 염기는 다음 그림 6과 같이 임의의 염색체에 추가되어 새로운 염색체를 생성하는 결과가 된다. 만약 중복되는 센서집합이 염기로 선택된 경우에는 축약을 통해 배제한다.

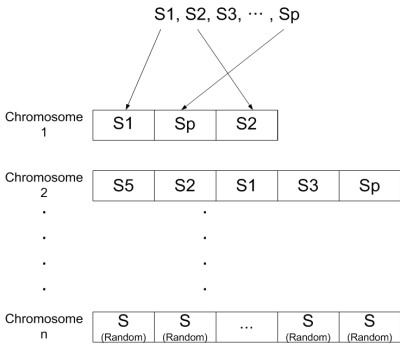

그림 6. 염색체의 부호화
Fig. 6 The Chromosome Encoding

하나의 센서가 모든 target을 모두 커버하는 경우에는 그 센서만을 요소로 갖는 하나의 집합이 된다. 하나의 염색체 인코딩과정을 무한정 반복할 수는 없으므로 임의의 시행횟수를 두어 그 동안에 구성된 집합들을 하나의 해로 간주하여 염색체로 설정한다. 제안한 인코딩 방법의 적절성을 Kobayashi 등이 제안한 코드화의 평가규범[17]에 따라 평가하였다.

코드화의 평가규범은 완비성(completeness), 건전성(soundness), 비중복성(non-redundancy)이다. 완비성은 문제 공간에서 해의 후보는 모두 염색체로 표현 가능하다는 것을 나타내는 것으로, 제안한 방법에서는 모든 target을 커버하는 센서 집합을 염기로 하여 염색체가 형성되므로, 모든 해의 후보가 염색체로 표현될 수 있다. 건전성은 유전자 알고리즘 공간의 염색체는 모두 문제 공간에서 해의 후보에 대응될 수 있음을 나타내는 것으로, 제안한 방법의 염색체는 그 크기를 노드와 target 수를 고려한 최대 길이로 구성할 경우에 모든 경우를 표현할 수 있으므로 건전성을 만족한다. 비중복성은 염색체와 해의 후보는 1 대 1로 대응될 수 있음을 나타내고, 제안한 방법에서는 하나의 염색체가 하나의 해 후보에 대응될 수 있다.

이상과 같이 제안한 방법은 완비성, 건전성, 비중복성을 모두 만족하므로 유전자 알고리즘에서 사용되는 인코딩방법으로 적절하다.

3.4.2 선택연산

제안한 유전자 알고리즘에서는, 최적해 수렴속도를 향상시키기 위한 선택연산 방법으로서 엘리트 보존 기법을 적용하였다. 또한, 각 염색체의 적합도를 기준으로 하여 룰렛-휠(roulette-wheel) 선택기법을 병행하여 사용함으로써 우수형질의 염색체 요소가 다음 세대에서도 살아남을 수 있는 확률을 높였다. 룰렛-휠 선택기법은 각 염색체의 적합도에 따라서 선택연산을 통해 선택될 확률이 비례적으로 할당된다. 즉, 각 염색체의 적합도를 모두 합한 값만큼의 크기를 가진 룰렛-휠을 가정한다. 각 염색체는 이 룰렛-휠 상에 자신의 적합도 만큼의 공간을 배정받는다. 본 연구에서 사용한 적합도함수는 최적화식에서 최대화하고자 하는 네트워크 수명값이다.

3.4.3 교차연산

IDS집합 문제에서 염색체의 각 염기들은 모든 target을 커버하는 노드의 집합이다. 이들의 조합은 다양하게 생성될 수 있으며 이러한 집합들을 구성하는 센서노드의 수도 다양하다. 따라서 유전자 알고리즘을 이용하여 IDS집합 문제를 해결할 때 염색체의 인코딩은 가변 길이를 가질 수 있도록 하여야 하고, 이를 제안한 방법에서는 구현하였다. 이와 더불어서 교차연산에서도 가변길이를 갖는 염색체들에 대한 교차연산이 수행될 수 있어야 하기 때문에 기존의 교차연산으로는 한계가 있어 새로운 교차연산이 요구된다. 본 연구에서 제안한 교차연산은 2단계의 수행과정으로 구성하였으며, 1단계에서는 교차연산의 계산시간을 단축시키기 위하여 가장 간단한 1점 교차를 수행한다. 즉, 2개의 부모염색체를 선택하여 각 부모염색체에 대해 임의의 교차점을 하나씩 선택한 후 그 교차점을 기준으로 교차점 이후의 염기들을 교환함으로써 교차를 수행한다. 2단계에서는 1단계의 결과로 구성된 염색체로부터 축양과정을 통하여 중복되는 염색체를 배재한다.

3.4.4 돌연변이연산

돌연변이 연산은 염색체를 구성함에 있어서 다양성과 가능한 넓은 해의 공간을 탐색하여 최적의 값을 찾기 위한 목적으로 사용된다. 본 연구에서 사용한 돌연변이 연산은 여러 개의 염기로 구성된 염색체에서 임의의 한 염기를 다른 염기와 교체함으로서 구현한다. 제안하는 유전자 알고리즘방식에서, 선택연산에 룰렛-휠 선택과 엘리트보존기법을 병용함으로써 최적해로의 수렴속도가 향상되는 반면 유전자 알고리즘의 다양성이 떨어질 수 있다. 이러한 단점을 보완할 수 있도록 돌연변이 연산을 설계한다.

IDS집합 문제에서, 유전자 알고리즘의 염색체의 길이가 길어진다는 의미는 노드의 집합수가 많다는 것을 의미하고 노드 집합의 수가 많을수록 노드 자체의 에너지를 나누어서 사용할 수 있으므로 전체 네트워크의 수명이 길어진다는 가정을 한다. 전형적인 돌연변이 연산을 IDS집합 문제에 적용하게 되면 염기의 중복이 발생할 수 있으므로, 이를 해결하고 염색체의 크기가 클수록 최적해 집합에 가까워질 확률이 높아지는 IDS집합 문제의 특성을 고려하여 새로운 돌연변이연산을 설계하고 제안한다.

본 논문에서 제안한 돌연변이연산은 그림 7과 같이 3단계의 수행과정을 가진다. 먼저 임의의 염색체 2개를 선택하여 두 염색체를 결합하는 결합과정을 통해 염색체의 크기를 키운다. 이는 IDS집합 문제에서는 한 센서가 더 많은 집합에 참여할수록 그 평가도가 높아지게 되므로, 염색체를 구성하는 염기의 수가 증가하게 되면 최적해에 근접하게 될 확률이 높아지게 된다.

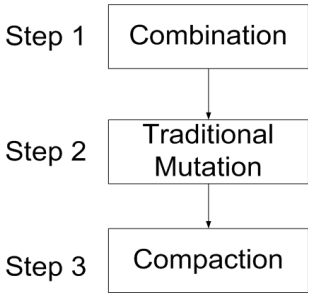

그림 7. 돌연변이연산의 과정
Fig. 7 The process of mutation

두 번째 단계에서 임의의 한 염기를 선택하여 다른 염기로 치환하는 전형적인 돌연변이연산을 적용함으로써 새로운 염색체가 생성된다(Step 2). 두 염색체를 결합하면 중복되는 염기가 발생할 수 있으므로 축약을 통해서 중복염색체를 제거한다(Step 3).

3.4.2 평가함수

유전자 알고리즘에서 임의의 염색체에 대한 적합도는 평가는 주어진 문제의 최적화 값에 의하여 결정된다. 염색체를 평가함에 있어서 그 염색체가 나타내는 수명값을 평가함수로 한다. 유전자 알고리즘에서 사용된 평가함수는 식(2) 과 같다.

$$f = \sum_{i=1}^{p} t_i \qquad (2)$$

여기서, t_i는 선택된 염기의 서비스 시간,즉 노드집합들의 동작시간이고, p는 염색체 수이다. 즉, 한 염색체를 구성하는 염기들의 집합에 대하여 동작시간을 설정하고 이 염기들의 동작시간 합산한 것이 선택된 염색체의 수명이다.

3.4.5 알고리즘의 동작

유전자 알고리즘의 실행순서는 다음과 같다.

단계 1: 염색체를 부호화하여 생성
단계 2: 초기 개체군을 생성
단계 3: 알고리즘을 수행하여 엘리트 선정
단계 4: 최종 세대수가 될 때까지 최고의 적합도를 만족하는 엘리트 선정

이상의 동작은 BC에서 운영되는 것을 고려하였으며 이러한 방식을 중앙처리 방식의 유전자 알고리즘(CGA : centralized genetic algorithm)라 한다.

Ⅳ. 시뮬레이션

전체 센서네트워크를 커버하는 새로운 구조의 IDS를 제안-모델링한 IDS 집합문제에 대한 해결방안으로 유전자 알고리즘 방법(여기서는 CGA 방법) 의 유효성 확인하기 위하여 시뮬레이션을 수행하였다. 노드가 분포는 지역은 500 x 500, 통신반경은 100m로 설정하였다. 시뮬레이션에 사용된 프로그램은 Java (SE version 1.6.0_24)로 구현하였으며, HP proliont ML350(INTEL zeon 3.0GHz, 4 CPU)에서 MS Server 2003 EE 운영체제에서 수행되었다.

먼저 본 연구에서 제안한 CGA에서 교차율에 따른 IDS 수명변화를 그림7에 나타내었다.

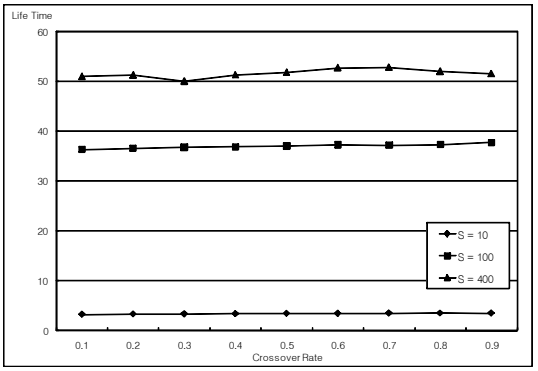

그림 8. CGA에서 교차율 변화에 따른 IDS 수명
Fig. 8 IDS lifetime on crossover rate in CGA

시뮬레이션 요소로서 유전자 알고리즘의 세대수는 100, 염색체 길이는 10, 그리고 교차율과 돌연변이율은 0.1부터 0.9까지의 변화를 주었으며 활동노드 수는 5개로 고정하였으며, 분포되는 센서의 노드 수는 10, 100, 400 개를 각각 적용시켰다. 그림 8에 나타나 반화 같이 노드의 수가 적을 경우에는 교차율의 변화에 반응이 미비하지만 노드의 수가 많은 경우 0.3이상의 값을 갖는 것이 바람직하다.

돌연변이율에 따른 IDS 수명 변화를 관찰한 결과는 그림 9와 같다. 그림 9에서 노드의 수가 적을 경우에는 돌연변이율의 변화에 민감하지 않지만 노드의 수가 많을 경우 0.3 이상에서 비교적 큰 변화를 볼 수 있다.

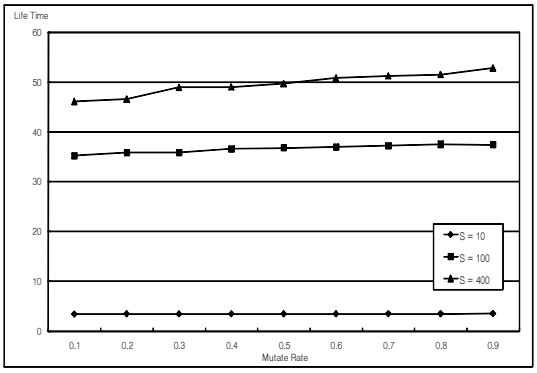

그림 9. CGA에서 돌연변이율 변화에 따른 IDS 수명
Fig. 9 IDS lifetime on mutation rate in CGA

돌연변이율이 높을수록 보다 많은 해의 공간을 탐색하는 것이 가능하므로 비교적 좋은 결과를 보이고 있으나 연산 량이 많아지는 단점이 있다. 제안한 유전자 알고리즘방식의 유효성 확인을 위하여 Greedy 방법을 구현하여 비교하였다. 그림 10에서는 노드의 수는 100 ~ 1400, 활동노드의수는 5로 고정하였다.

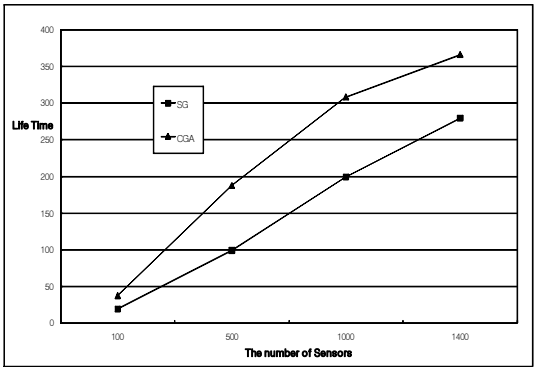

그림 10. 활동 센서노드 수에 따른 IDS 수명비교
Fig. 10 The comparison of IDS life as lifetime as active node number

3.2절에서 언급한 Greedy 방법에서 disjoint 조건에서 해를 구한 것(SG : Simple Greedy)과 CGA에서는 그림 8, 9에서 나타난바와 같이 교차율 0.9, 돌연변이율은 0.7이면 충분한 효과가 있다고 판단되어 설정하였으며 세대수는 100으로 하였다. 결과 disjoint 조건에서 Greedy 방법보다 유전자 알고리즘 방식이 보다 우수한 결과를 보였다.

Ⅴ. 결 론

본 연구에서는 센서네트워크에서 모든 패킷을 관찰할 수 있는 IDS 구축방법을 제안하고 이를 모델링하였으며, 제안된 최적화 식에 대하여 에너지 효율성을 제고하는 방안을 제안하였다. 유전자알고리즘에 의한 방식의 해를 제안하였으며 Greedy 알고리즘과 비교하여 제안한 방법과 모델의 유효성을 확인하였다. Greedy 방법은 계산 량이 비교적 작다는 장점이 있으나 IDS에서 연속적인 관찰이라는 QoS에 문제가 있다. 유전자 알고리즘 방법은 상대적으로 계산 량은 많으나 최적의 IDS 집합을 한 번의 계산으로 구하기 때문에 연속적인 관찰이 가능해진다. 향후 연구과제로서는 가변의 통신반경 기능을 이용한 에너지 효율성을 고려한 IDS 스케줄링 기법, 운영모드 변환에 따른 context 스위칭을 고려한 운영기법 등이다.

참고문헌

[1] H. Chan and A. Perrig, “”Security and Privacy in Sensor Networks,”” IEEE Comp. Mag., pp. 103 – 105. Oct. 2003.

[2] E. Shi and A. Perrig, “Designing Secure Sensor Networks,” Wireless Commun. Mag., vol. 11, no. 6, pp. 38 – 43, Dec. 2004.

[3] I. F. Akyildiz et al., “”A Survey on Sensor Networks,”” IEEE Commun. Mag., vol. 40, no. 8, pp. 102 – 114., Aug. 2002.

[4] E. Shi; A. Peerig, ″Designing Secure Sensor Networks″, Wireless Communications Magazine., vol. 11, no. 6, 2004.

[5] Yang Xiao “Security in Distributed, GRID, and Pervasive Computing″, Chapter 17(Wireless Sensor Network Security: A Survey), CRC Press, 2006.
[6] Y. Wang; G. Attebury; B. Ramamurthy, ″ A Survey of Security Issues in Wireless Sensor Networks, ″IEEE Comm. Surveys & Tutorials, vol. 8, no. 2, 2006.
[7] I. Sato, Y. Okazaki, and S. Goto. An improved intrusion detection method based on process profiling. IPSJ Journal, 43(11):3316 – 3326, 2002.
[8] Edith C.H. Ngai, “Intrusion Detection for wireless Sensor Networks”, The Chinese University of Hong Kong Department of Computer Science and Engineering Ph.D. Term 2 paper, pp. 29-37, 2005.
[9] R. Roman; J. Zhou; J. Lopez, ″Applying Intrusion Detection Systems to Wireless SensorNetworks,″ in Proc. CCNC’06, 2006.
[10] F. Anjum; D. Subhadrabandhu; S. Sarkar; R. Shetty, ″On Optimal Placement of Intrusion Detection Modules in Wireless Sensor Networks,” BROADNETS’04, 2004.
[11] A. Agah; S. Das; K. Basu, ″Intrusion Detection in Sensor Networks: A non-cooperative Game Approach,″ IEEE ISNCA, 2004.
[12] C. Loo; M. Ng; C. Leckie; M. Palaniswami, ″Intrusion Detection for Routing Attacks in Sensor Networks,″ Int’l. Journal of Distr. Sensor Networks, vol. 2, 2006.
[13] C. Su; K. Chang; Y. Kuo; M. Horng, ″The New Intrusion Prevention and Detection Approaches for Clustering-Based Sensor Networks,” IEEE WCNC, 2005.
[14] MICA2 Radio Stack for TinyOS. http://www.tinyos.net/tinyos-1.x/doc/mica2rad, 2007.
[15] Techateerawat, P. and Jennings, A. ″Energy Efficiency of Intrusion Detection Systems in Wireless Sensor Networks″, IEEE/WIC/ACM International Conference on Web Intelligence and International Agent Technology Workshops, 2006.
[16] M. Cardei, D.-Z. Du, Improving Wireless Sensor Network Lifetime through Power Aware Organization, ACM Wireless Networks, Vol 11, No 3, May 2005.
[17] 문병로, 유전알고리즘, 두양사, 2003.

저자소개

성기택(Ki-taek Seong)

1988, 1991: 부경대학교 전자통신공학과 학사, 공학석사
2007: 부경대학교 정보시스템학과 공학박사

1992년 ~ 1997년 국방과학연구소 선임연구원
현재 동명대학교 정보보호학과 (조교수)
※관심분야 : 네트워크 보안, 센서네트워크