해양 USN 환경에서 수질환경의 온라인 정상·비정상 상태 구분

정신출* · 정희택**

Online Identification for Normal and Abnormal Status of Water Quality on Ocean USN

Sin-Chul Jeoung* · Hee-Taek Ceong**

요 약

본 논문은 해양 USN 환경에서 수질환경의 온라인 정상·비정상 상태를 구분하기 위한 기법을 제안한다. 해양 수질 환경 인자의 정상 상태 집합을 정의하기 위해 정상·비정상 구분 특성을 갖는 인공면역시스템의 부정선택 알고리즘을 활용한다. 비정상 상태 구분을 위해 해양 USN 환경에서 센서를 통해 수집된 해양 수질 환경의 정상 집합 생성이 필요하다. 이를 위해 각 측정 인자에 대한 단위 데이터의 돌연변이와 해양 수질 상태를 각 요소의 논리곱적 관점에서 상태 데이터의 돌연변이를 기반으로 정상 집합을 생성한다. 생성된 정상 집합을 활용하여 비정상 상태를 구분한다. 이 과정을 가우시안 함수를 기반으로 돌연변이 된 정상 집합에 대하여 모의 실험을 제시한다. 이렇게 제안된 기법을 해양 수질 센서 로거단에 설치함으로써, 온라인으로 해양 수질 환경의 정상 인자를 모니터링 할 수 있다.

ABSTRACT

This paper suggests the online method to identify normal and abnormal state of water quality on the ocean USN. To define normal of the ocean water quality, we utilize the negative selection algorithm of artificial immunity system which has self and nonself identification characteristics. To distinguish abnormal status, normal state set of the ocean water quality needs to be defined. For this purpose, we generate normal state set base on mutations of each data and mutation of the data as logical product. This mutated normal (or self) sets used to identify abnormal status of the water quality. We represent the experimental result about mutated self set with the Gaussian function. Through setting the method on the ocean sensor logger, we can monitor whether the ocean water quality is normal or abnormal state by online.

키워드

AIS(Artificial Immune System), Negative Selection, USN(Ubiquitous Sensor network), Mutation, Gaussian Function, 인공면역시스템, 부정선택, USN, 돌연변이, 가우시안 함수

Ⅰ. 서 론

유비쿼터스 기술이 발전하면서 핵심 기술인 USN (Ubiquitous Sensor network)에 대한 관심이 높아지고 있다. USN은 태그와 센서[1]로부터 사물이나 환경 정보를 센서링하거나 이를 모니터링하여 유용한 최종

* 전남대학교 디지털컨버전스(msjjbbr@nate.com)
** 교신저자 : 전남대학교 교수(htceong@jnu.ac.kr
접수일자 : 2012. 07. 10 심사(수정일자) : 2012. 07. 26 게재확정일자 : 2012. 08. 09

정보로 활용하고 신속하게 관리하는 지능형 사회기반 인프라이다[2, 3, 4]. 언제, 어디서, 어떠한 기기를 이용하더라도 자유롭게 접속해 필요한 정보를 컴퓨팅 할 수 있는 융합적인 환경을 구현함으로써 실세계 문제들에 대한 해결책으로 적용되고 있다. 특히 해양 USN 환경에서는 해양변화의 모니터링을 기반으로 한다. 해양의 환경을 모니터링하기 위해 용존산소, 염도, 또는 해수온도 등 데이터를 수집하고, 이 데이터를 분석하는 것이 주요 핵심이다. 청수대, 냉수대와 같이 급격한 해양 환경 변화에 대해 효과적으로 대응하기 위해, 해양 USN 환경에서 수집된 정보가 정상 인자인지 비정상 인자인지 수질환경 인자의 변화를 구별하는 것은 매우 중요하다.

해양 수질 환경 인자의 정상·비정상 문제는 인공면역시스템(Artificial Immune System)의 기법을 적용할 수 있다. 인공면역시스템은 자연계의 면역시스템을 기계적으로 모사한 방식이다. 사람의 몸은 끊임없이 외부의 공격을 받으며 살아간다. 때문에 외부 물질 즉 항원(Antigen)으로부터 우리의 신체를 방어하기 위해 정상·비정상 인자를 구별 할 수 있는 메카니즘이 중요하다. 태어나면서 이미 존재하는 면역성을 선천적 면역시스템이라고 한다면 선천적으로 갖고 있지 않은 다른 물질이 발견되면 이를 제거하거나 효과적인 대응을 하는 과정을 적응적 면역시스템이라 부른다. 적응적 면역시스템은 학습, 적응, 기억 능력 등 중요한 특성을 기반으로 정상·비정상을 구별하는 능력이 있다. 이러한 특성은 해양 USN 환경에서 정상·비정상 탐지와 동일한 특성이 존재한다. 즉, 사람의 적응적 면역시스템을 모사한 인공면역시스템 원리를 해양 USN 환경에서 적용하여, 해양 수질 환경의 정상·비정상을 식별 할 수 있다.

본 연구에서는 인공면역시스템 기법을 활용하여 해양 USN 환경의 해양 수질 환경 인자의 정상·비정상 상태를 구별하는 기법을 제안 한다. 정상 및 비정상 상태를 구별하는 것은 인공 면역시스템의 부정선택 기법을 활용한다. 비정상에 대한 식별을 하기 위해, 해양 USN 환경에서 대량의 해양 수질 환경 데이터는 특정 상태를 의미하는 상수화된 점(Point) 정보이다. 때문에 해양 수질 환경 인자의 정상·비정상 상태를 구별하기위해 자기 상태를 정확하게 반영할 수 있는 데이터 돌연변이의 자가 집합을 생성하고 이를 활용하여 해양 수질 환경 인자에 대해 모의실험을 수행한다. 따라서 본 논문에서는 인공면역시스템의 적응적 면역기법과 돌연변이 확장을 통하여 급격한 변화에 따른 해양 수질 환경 인자의 비정상을 효과적으로 탐지 할 수 있으며 해양 USN 환경의 온라인 시스템에서 탐지가 가능하다.

본 논문의 구성은 다음과 같다. 2장에서는 인공면역시스템과 이를 활용한 관련 연구들을 정리하고 3장에서는 해양 USN 환경에서 인공면역시스템을 활용한 자기(Self)와 비자기(Non-self)를 구별하기위해 해양 수질 환경을 조사 분석 하고 해양 수질 환경의 자기, 비자기 탐지 방안을 제안하고 4장에서는 이를 보이기 위해 실험을 한다. 마지막으로 5장에서는 결론 및 향후 연구에 대하여 기술한다.

II. 관련연구

2.1 인공면역시스템

생물학적 면역시스템은 항원으로 통칭하는 다양한 외부 물질이나 단백질에 대하여 신체를 방어할 수 있는 정교하고 다양한 구조로써 자기 자신과 외부 물질을 구분하는 탁월한 능력을 가지고 있다. 이러한 자연계의 면역시스템으로부터 영향을 받아 컴퓨터상에서 해결해야 할 문제에 대한 새로운 접근법을 인공면역시스템이라고 한다. 인공면역시스템은 다양한 응용분야에서 문제들에 대한 해결 방안들을 제공하고 있다. 본 연구에서도 해양 수질 환경을 하나의 거대한 면역체계로 보고, 인공면역시스템의 기법을 활용한다.

태어날 때부터 이미 면역성이 존재하는 것이 선천적 면역시스템이라면 선천적으로 갖고 있지 않은 다른 물질에 대응하기 위한 여러 가지 방법과 적용 과정은 적응적 면역시스템이다. 선천적 면역시스템은 외부 항원에 대해 즉각적인 반응이 가능하기 때문에 1차 항원에 대한 제거가 빠르지만 특이성을 구분하는 능력이 떨어짐으로 적응적 면역시스템을 통해 특이성을 구분한다[5]. 이 처리 과정의 기본적인 두 가지 요소는 B 세포와 T 세포로 매우 중요한 역할을 수행한다. B 세포는 항원들과 연결 할 수 있는 특정한 단백질이 존재한다. B 세포는 특정 침입 유기체 즉, 항원을 발견하고 파괴하기 위한 다양한 항체를 생성하

며 특정 항원에 신속하고 효율적으로 반응 할 수 있는 세포로 복제된다. 또한 일부 항체는 항원과의 적합성을 높이기 위해 돌연변이 된다. 이때 T 세포는 도움 T 세포(Helper T-cell)에 의해 B 세포의 분화와 항체생산을 도와 면역반응을 촉진시켜 주며 억제성 T 세포(Superessor T-cell)는 항체의 활동을 억제하기도 한다. 또한 킬러 T 세포(Killer T-cell)를 통해 해로운 세포를 파괴하여 면역기능을 강화해준다. 모든 항원에 대해 즉각적으로 반응하는 선천적 면역시스템에 비해 적응적 면역시스템은 알려지지 않았던 외부 물질에 대한 방어체계를 준비하는데 일정시간이 필요하지만 기억능력 즉 학습된 기억을 가지고 있기 때문에 외부 물질이 2차적인 침투에 대한 방어능력이 뛰어나다.

일반적으로 자신의 정상 세포를 자기라 하고 외부 물질을 비자기라고 할 때 효과적인 방어 체계를 구축하기 위해 자기 세포에 해당하지 않는 비자기의 구별이 필수적이다. 아래의 그림 1은 T 세포와 B 세포의 항원 인식 및 활성화 메카니즘을 나타낸 것이다. 적응적 면역시스템에서 항원 인식 과정은 크게 T 세포, B 세포, 항원전달세포(Antigen-presenting cell)에 의해 이루어진다[6, 7]. (Ⅰ) T 세포는 항원을 직접 인식하지 못하므로 항원전달세포에 의해 항원을 인식하는 과정이다. (Ⅱ) 항원전달세포가 항원을 섭취하고 단백질(Peptide) 조각들로 이루어진 항원은 합성된 단백질 조각을 그 세포의 표면에 보여주는 주조직 적합 복합체(Major Histocompatibility Complex) 분자와 반응하여 항원전달세포의 표면에 주조직 적합 복합체와 단백질 복합체로 발생한다. (Ⅲ)T 세포가 항원전달세포를 인식 하면 (Ⅳ)활성화 된 T 세포가 특정 단백질인 림포카인(lymphokine)을 분비함으로써 면역 반응을 활성화 시키며 T 세포에 의해 B 세포도 활성화가 된다. B 세포는 주조직 적합 복합체 분자의 도움 없이 항원을 인식할 수 있다. (Ⅴ)단계에서 수용체를 이용하여 특정 항원에 직접 반응한다. 항원 결합 반응이 일어나면 B 세포는 활성화 되어 급격히 증식하게 되고 (Ⅵ)단계에서 항체를 생산하는 플라즈마 세포(plasma cell)가 되거나 일부 기억세포(memory cell)가 된다. (Ⅶ)플라즈마 세포는 항원을 파괴하고 기억세포는 항원이 소멸하더라도 사라지지 않고 항원에 대한 면역기능을 유지시킨다. 따라서 같은 항원이 재 침투했을 때 감염에 대한 학습 능력으로 인해 선천적 면역시스템 보다 효과적으로 항원에 대한 면역시스템을 재 작동시킨다.

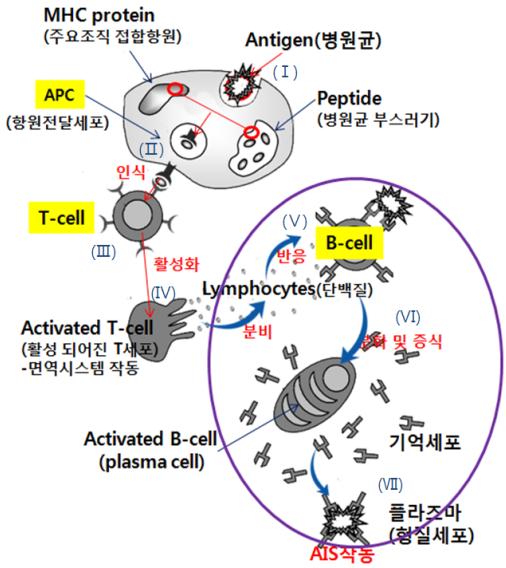

그림 1. 항원 인식 및 활성화 메카니즘
Fig. 1 Antigen recognition and activation mechanism

항원은 자기 항원과 외부 병원균으로 나눌 수 있다. 우리 몸을 보호하고 외부 물질을 파괴하기 위해 프로그램 되어 있는 항체는 항원에 반응할 수 있는 매우 다양한 형태로 만들어져 있다. 그러나 자기를 공격하는 돌연변이가 자기인지 비자기인지 구별하기 어렵기 때문에 자기질병 즉 면역병을 일으킬 수 있다. 림프구를 통해 항체를 생성할 때 T 세포는 B 세포에게 다양한 정보를 제공하여 항체생성에 도움을 주고 B 세포의 다양한 세포분화로 항체가 증식된다. 그러나 자기를 항원으로 인식하는 것을 방지하기 위해 T 세포는 자가 면역성을 가진 T 세포의 면역반응을 촉진하는 것이 아니라 오히려 자가 면역성을 가진 T 세포를 제거하고 자기요소와 반응하지 않는 세포를 선택하여 자가 면역 반응을 차단한다. 이것을 부정선택(Negative selection)이라고 한다.

부정선택 알고리즘은 림프구에서 T 세포들의 성숙과정에 의해 영향 받은 것이다. 그 알고리즘은 검열(censoring)과 감시(monitoring) 단계로 구성된다. 검열

단계는 변화 – 자기·비자기 집합(change-detectors)을 준비한다. 감시 단계에서, 검열단계에서 생성된 자기·비자기 집합을 사용하여 변화를 감시 한다. 확장하여 생성된 자기집합과 새로운 자기가 일치 되면 모니터링 과정에서 비정상으로 추출되어 버려지고 일치하지 않는다면 유용한 자기집합에 추가된다.

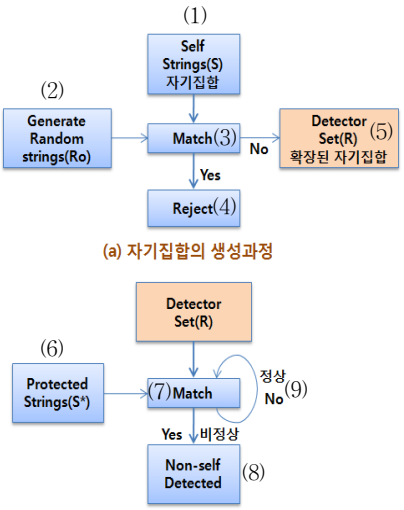

그림 2. 부정선택 알고리즘과 모니터링
Fig. 2 The negative selection algorithm and monitoring

포괄적인 자기 및 비자기 집합 생성 알고리즘은 [8]에 의해 제안된 기법으로 그 알고리즘은 자기 자신과 반응하는 T 림프구(T-lymphocytes)의 제거 과정과 같다. (1)자기집합(S)을 정의하고 (2)하나의 임의의 후보 탐지기(Ro)를 생성하여 (3)각각 생성된 후보탐지기(Ro)와 자기집합(S)을 일치 시킨다. (4)만약 생성된 후보탐지기(Ro)와 자기집합(S)이 일치한다면 그 후보 탐지기는 제거되지만 (5)생성된 후보탐지기(Ro)와 자기집합(S)이 일치하지 않는다면 그 후보 탐지기는 유용한 탐지기(R)의 집합에 추가된다. 이러한 과정을 반복적으로 수행하여 탐지기 집합(R)과 (6)새로운 데이터(Protected strings(S*)를 (7)일치시켰을 때 입력 값이 탐지기 집합(R)과 일치가 되면 (8)비정상 이라고 하고 새로운 데이터와 일치 시켜서 입력 값이 탐지기 집합(R)과 일치가 안 되면 (9)정상 이라고 감시됨에 따라 임의의 입력 값이나 자기 자신의 변화가 발생할 때 마다 정의된 항원들과의 유사성을 계산함으로써 이 값이 자기인지 비자기인지 구별 할 수 있다. 이 알고리즘은 그림 2와 같다.

2. 2 인공면역시스템을 활용한 연구들

[9]에서는 인공 면역시스템과 분산 유전자 알고리즘에 기반하여 동작하는 자율 분산 로봇 시스템에서 로봇들의 협조행동을 학습하는 방법을 제안하였다. 자율분산시스템은 자율분산제어방식을 통해 다수의 로봇을 제어하는 시스템으로 다른 로봇들과의 네트워크를 통해 유기적인 협동을 수행하기 위한 각 로봇의 역할을 결정하게 된다. 그리고 각기 역할을 수행하는 로봇들이 모여 전체적인 시스템을 구성하게 된다. 각 로봇들의 협조 행동에 관한 문제는 면역시스템의 T 세포와 B 세포의 상호 커뮤니케이션을 통하여 유전자의 재결합이나 돌연변이로 인한 항체 생산 등 면역시스템의 정보처리 메카니즘을 모델링하였다. 그리고 각 로봇들이 환경변화에 대응하여 작업의 효율성을 높이기 위한 학습 알고리즘으로 분산 유전자 알고리즘을 확장함에 따라서 인공 면역네트워크로 작업을 탐색하는데 빠르게 수행하기 위한 전역행동과 탐색된 작업에 대한 로봇들의 협조행동에 대한 학습 결과를 수행하는 지역행동의 상호작용으로 적응성이 향상된 로봇 자율분산시스템을 구성하였다. 이러한 시스템을 구현하기 위해서 전역행동에서의 매우 중요한 문제는 주변에서 탐색된 작업과 로봇들을 모으는 과정이다. 이를 항원으로 하고 전역행동을 항체로 하여 면역네트워크 방정식을 적용하는 반면 해양 USN 환경에서는 전역행동과 같은 역할을 부정선택기법을 사용하였으며 이를 수행하는 지역행동의 역할로서 자가 집합인 학습데이터를 기반으로 가우시안 돌연변이 기법을 적용하였다. 그러므로 다양한 해양환경의 변화에 대처하기 위한 데이터 돌연변이를 통해 변화에 유연하고 강한 정상 인자를 구별할 수 있다.

[10]에서는 고장허용(fault-tolerant) 하드웨어에서 가장 기본이 되는 온라인 하드웨어 테스트 시스템 구현을 위하여 자동 고장 검출 기능을 가진 알고리즘이 제안되었다. 이 알고리즘은 면역시스템에서의 항체의 다양성(antibody diversity) 원리와 항원전달세포를 이

용하여 허용조건을 생성하고 생성된 허용조건으로 자기와 비자기를 구분하였다. 먼저 항원전달세포가 허용조건이 정상인지 비정상인지 사전 제시를 하고 그로 인해 자기를 검출한다. 그리고 부정선택을 이용하여 허용조건 집합을 생성하여 비자기 발생 시 에러 검출을 한다. 이때 1차적으로 검출이 안 된 세포들은 부정선택과 항체의 다양성 원리에 의해 새롭게 생성된 허용조건을 통해 비자기로 구분되는 장점을 가지고 있다. 또한 시스템의 한계를 극복하기 위해 허용조건의 수를 줄임으로서 효율적인 하드웨어 시스템 구현이 가능하며 비자기 탐색율도 높일 수 있다. 또한 더 많은 비자기 탐지를 위해 항체 생성을 항원전달세포 콰인-맥클러스키(Quine-McClus key)방식으로 구현함으로써 항체를 가장 최적으로 생성할 수 있다는 연구결과를 발표하였다. 이는 해양 USN 환경에서 해양 수질 환경 인자를 정상에 포함시키기 위해 데이터 돌연변이의 자가 집합 생성과정과 유사하다. 항원전달세포와 항체의 다양성 원리는 부정선택 알고리즘을 활용하기위한 효과적인 조건을 통해 본 연구에서는 데이터 돌연변이를 부정선택한 후 자기를 검출해서 생성된 자가 집합과 임의의 데이터를 비교하여 조건에 맞는 새로운 탐지기 집합을 생성할 수 있다.

[11]에서는 인공면역시스템에 대한 이론을 중심으로 인공면역시스템의 복제선택, 기억세포, 부정선택, 수용체 조작, 백신, 면역네트워크 등 새로운 최적화 기법을 비교 분석 하였다. 인공면역시스템의 최적화 분야에 적용된 논문들에 대한 문헌 조사 결과 인공면역시스템의 해법으로 효과성, 유일성, 우수성, 수행 속도가 있다. 해의 우수성에 대하여는 부정적 선택이나 복제선택 기법의 경우 명백하게 우수하다고 판정하지 못하였으며, 면역네트워크에 대하여는 연산속도가 우수한 것으로 보고된바 최적화 문제에서 적합한 인공면역시스템을 설계하는 것이 중요하다고 판단하였다. 또한 면역시스템의 비교 분석한 논문 등을 혼성한 논문들은 각 분야의 장점을 공유하기 위한 것으로서 새로운 생물학적 연구 결과가 인공면역시스템의 효과성을 높여줄 것이라고 제시하였다. 따라서 인공면역시스템의 면역네트워크는 분산적 의사 결정 구조를 가지고 있으므로 분산적인 시스템 전체의 효율을 최적화해야하는 문제에 이용한다면 효과가 있을 것으로 판단된다고 제시하였다.

[5]에서는 인공면역시스템을 활용한 최적화 알고리즘에 대한 새로운 방법을 제안하였다. 친화도 측정(affinity measurement)방식과 매칭 함수(matching function)를 통해 값을 측정할 때 해밍 거리(Hamming distance), 다중 인접 비트 규칙(Multiple contiguous bit rule), 가중치 거리(weighed distance) 세 가지 방식으로 살펴본 결과 다중 인접 비트 규칙의 경우가 인공 면역기반 최적화 알고리즘 중 가장 좋은 값을 가짐을 제시하였다. 또한 해밍 거리와 다중 인접 비트 규칙의 경우 항체-항원 간의 매치 횟수가 많다고 최적해가 좋아지는 것은 아님을 제시하였다.

[5, 11]에서 제시된, 인공면역시스템을 활용한 최적의 해를 찾는 연구들은 본 연구에서 대상으로 하는 해양 수질 환경의 자기 및 비자기 구별 문제와 다르지만, 활용하는 부정적 선택기법 및 돌연변이 기법의 관점에서 분석하였고 [5, 9, 10, 11]에서 선행된 연구들은, 본 연구에서 활용하고 있는 인공면역시스템을 이용하여 해당 문제를 접근하는 공통점이 있으나, 본 연구에서 분석 대상이 되는 수치화된 데이터에 대한 돌연변이를 고려하고 있지 않다.

III. 해양 USN환경에서 부정선택 기법을 활용한 자기와 비자기 구분

3.1 해양 수질 환경 데이터의 정상 상태

T 세포는 발생도중 유전자 재배열 과정을 통해 돌이변이가 만들어 진다. 이때 자기 단백질에 반응하는 T 세포는 파괴되며 반응하지 않는 T 세포는 몸을 순환하며 면역기능을 수행한다. 결국 비 일치된 세포들은 최종 집합인 자기가 되고 자기 집합을 이용하여 새로 들어오는 데이터를 모니터링하고 비자기 집합과 유사하다고 판단될 경우 데이터는 비자기로 구별할 수 있다[6]. 따라서 해양 수질 환경 정보의 정상상태를 설정하기 위해 인공면역 시스템에서 자가 및 비자가 구별 기법을 활용한다. 부정선택에서 활용하는 방법으로 새로이 생성된 데이터가 자가 집합에 속하지 않는 경우 새로운 자가 집합에 포함 시키는 기법이다. 본 연구에서는 자가 집합인 학습 데이터를 기반으로 돌연변이를 활용하여 확장한다. 이는 현재 자가 상태의 정보를 활용하면서 타당한 값을 새로운 자가 상태의 원소로 정의하기 위함이다. 이를 위해 해양 수질

환경의 단위 데이터에 대한 돌연 변이와 이들 요소들이 논리곱적 특성을 반영한 돌연변이 방법과 이를 통한 해양 수질 환경 데이터의 정상 상태 집합을 정의한다.

먼저, 해양 수질 환경에서 단위 데이터 (즉 수온, 염도, 용존산소, pH) 각각에 대한 돌연변이를 정의하기 위해 다음과 같이 단위 데이터 돌연변이를 정의한다.

<정의 1> 단위 데이터 돌연변이(Element Data Mutation)

해양 수질 환경 인자 Oe에 대해 다음을 만족하는 Oe'를 돌연변이 된 데이터라 정의한다.

$$Oe - \Psi N_{Oe}(\mu,\sigma) \le Oe' \le Oe + \Psi N_{Oe}(\mu,\sigma) \quad (1)$$

여기서 $Oe \in \{t,s,d,p\}$이며 t는 수온, s는 염도, d는 용존산소, p는 pH를 의미하고 $N_{Oe}(\mu,\sigma)$는 해당 수질 인자에 대한 가우시안 함수를 의미하며, Ψ는 돌연변이 비율을 의미한다.

한편, 해양 수질 상태는 각 단위 데이터의 개별적인 값에 의해 특성을 반영할 수 있지만, 특정 상태는 여러 인자들의 논리곱적 의미를 갖는다. 이를 반영하는 돌연변이를 정의하기 위해 다음과 같이 상태 데이터 돌연변이를 정의한다.

<정의 2> 상태 데이터 돌연변이(State Data Mutation)

해양 수질 상태 SOe에 대해 다음을 만족하는SOe'를 해양 수질 상태 데이터의 돌연변이라 정의한다.

$$SOe - \Psi N_{SOe}(\mu,\sigma) \le SOe' \le SOe - \Psi N_{SOe}(\mu,\sigma) \quad (2)$$

여기서 $SOe = t \times s \times d \times p$이며, $N_{SOe}(\mu,\sigma)$는 SOe에 대한 가우시안 함수를 의미하며, Ψ는 돌연변이 비율을 의미한다.

본 연구에서 단위 데이터 및 상태 데이터의 돌연변이는 범위로의 확장을 의미한다. 이는 공간상에서 하나의 점을 영역으로 확장함을 의미한다. 이는 하나의 점에서 다른 점으로 돌연변이 하는 기존 방법은 너무 작은 자기 집합(Self set)을 산출하게 되어 정상 상태임에도 비정상 상태로 결정하는 거짓 부정(false negative)이 발생하기 때문이다. 또한 해양 수질 환경은 수질 통계논문[12]에서 분석한 바와 같이 일정 값들 내에서의 변화를 갖기 때문에 범위로의 돌연변이를 수행한다. 또한 범위로의 돌연변이는 기존 자가 집합을 내포하고 있기 때문에 기존 자가 집합에 돌연변이 된 정보를 반영할 필요 없이, 최종 돌연변이 된 공간들의 집합만을 자가 집합으로 정의할 수 있다.

정의된 단위 데이터의 돌연변이와 상태 데이터 돌연변이에 의한 자가 집합을 다음과 같이 정의한다. 단위 데이터 및 상태 데이터 각각은 각 요소인자에 대한 자가 집합이냐 논리곱에 의한 자가 집합이냐를 구별하는 것이며 다음 장에서 제시하는 실험에서 구별하여 실험한다. 이는 의미적으로 동일한 상태를 단위 요소들의 합으로 보느냐 곱으로 보느냐의 의미를 갖는다.

<정의 3> 해양 수질 환경 데이터의 자기 집합

각 수질인자에 대한 데이터 돌연변이에 의해 생성된 데이터 집합은 해양수질 환경 데이터의 자가 집합으로 다음과 같이 정의한다.

$$\bigcup_{Oe' \in \{t,s,d,p\}} Oe' + \cup\, SOe' \quad (3)$$

3.2 해양 수질 데이터의 자가 집합을 활용한 해양 환경 모니터링

생성한 자가 집합을 기반으로 다음과 같은 해양 환경 모니터링을 수행한다. 먼저 단위 데이터의 돌연변이에 의한 자가 집합인 $\bigcup_{Oe' \in \{t,s,d,p\}} Oe'$에 대한 해양 환경 모니터링은 다음 그림과 같다. 각각 입력된 현재 수질 환경 데이터 각각에 대한 해당 인자의 자가 집합과 비교하여 자가 집합내에 속하는지를 판단한다. 각 요소가 해당 자가 집합에 속하면 현재 해양 수질 상태는 정상으로 판정한다. 만약 하나라도 해당 자가 집합에 속하지 않으면 비정상상태로 판별한다. 필요에 따라 특정 인자가 해당 자가 집합에 속하지 않을 경우에도 정상상태로 응용해서 정의할 수 있다. 예로 들어 특정 센서가 고장 나거나 없는 경우에 허용할 수 있다.

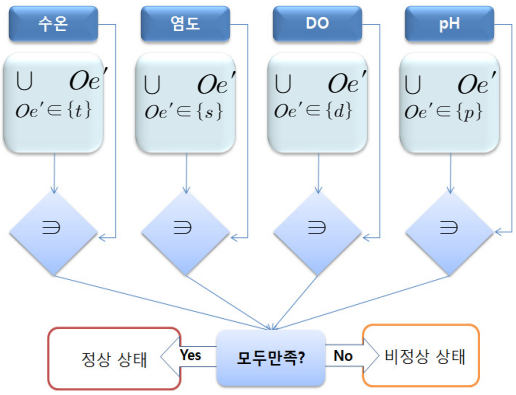

그림 3. 단위 데이터 돌연변이로 구성된 자가 집합을 활용한 해양환경 모니터링
Fig. 3 Ocean environment monitoring using self set the mutation of each data

다음으로 상태 데이터 돌연변이에 의해 구성된 자가 집합에 대한 해양 환경 모니터링은 다음과 같다. 모든 인자에 대한 논리곱을 통해 자가 집합과 비교하여 자가 집합에 속하면 정상상태로 식별한다.

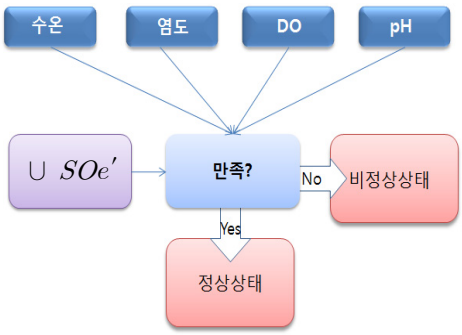

그림 4. 상태 데이터 돌연변이로 구성된 자가 집합을 활용한 해양환경 모니터링
Fig. 4 Ocean environment monitoring using self set status data mutation

V. 실 험

해양 USN 환경에서 측정된 데이터를 기반으로 실험을 수행하였다. 고흥 지역에 YSI사의 센서[13]로 장기간 안정적인 해양 환경 측정을 위해, 광기반의 용존산소 측정기를 포함하고 있다. 측정된 데이터는 2009년에서 2010년 데이터이며 측정 정보는 수온, 염도, 용존산소(DO), pH이다. 측정은 매분 측정되었으나 1일 변화가 크지 않기 때문에 본 연구에서는 측정된 값의 평균을 산출하여 그 날의 측정치로 하였다. 또한 일반성에 위배됨이 없이 2009년과 2010년에 측정된 데이터는 특이한 해양 환경 변화가 없다고 판단하고 해양의 정상 상태를 의미한다고 가정하였다. 한편, 센서의 관리 과정에서 측정된 이상 데이터는 실험을 위해 제외하였다. 인공면역 시스템을 활용한 해양 환경의 정상 및 비정상 구분을 위해 본 연구에서는 다음과 같은 데이터를 학습과 검증을 위한 데이터로 사용한다.

표 1. 실험 데이터
Table 1. The test data

년도	데이터 개수	비고
2009년	285개	학습 데이터
2010년	253개	검증 데이터

4.1 해양 수질 데이터의 확률분포

단위 데이터 돌연변이를 위한 수질 데이터들의 확률분포

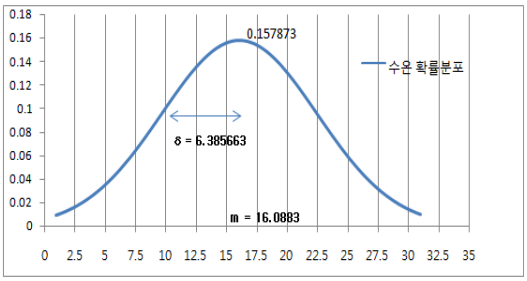

그림 5. 수온의 확률분포
Fig. 5 Probability distribution of water temperature

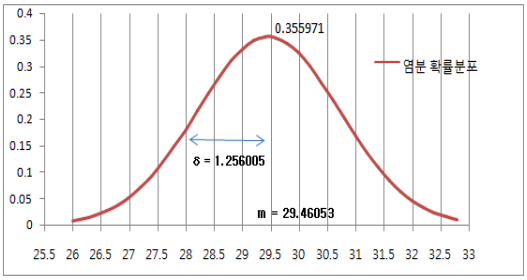

그림 6. 염도의 확률분포
Fig. 6 Probability distribution of salinity

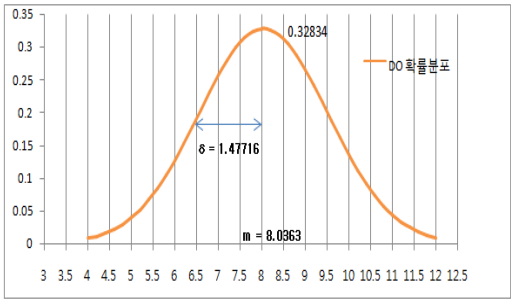

그림 7. DO의 확률분포
Fig. 7 Probability distribution of DO

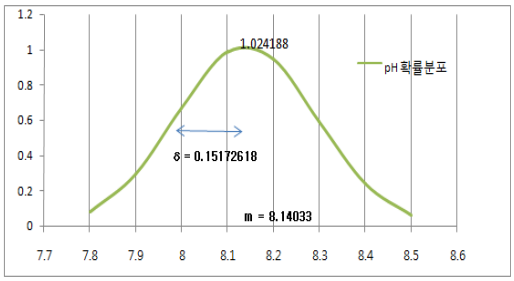

그림 8. pH의 확률분포
Fig. 8 Probability distribution of pH

해양수질 환경 데이터의 자기 집합을 생성하기 위해, 먼저 단위 데이터 돌연변이를 생성해야한다. 가우시안 함수를 기반으로 한 돌연변이를 위해 다음과 같이 각 데이터 인자에 대한 확률분포를 분석하였다.

다음으로 상태 데이터 돌연변이를 위한 수질 인자들의 논리곱에 대한 가우시안 함수를 생성하기 위해 다음과 같은 확률분포를 생성하였다.

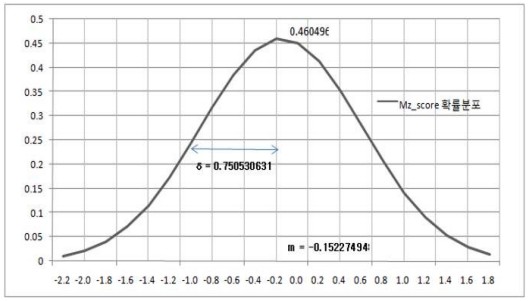

그림 9. 수질 데이터들의 논리곱에 대한 확률분포
Fig. 9 Probability distribution about of data as logical product

4.2 단위 데이터 돌연변이의 자가 집합에 대한 해양환경 모니터링 실험

2009년 데이터를 기반으로 단위 데이터 돌연변이의 자가 집합을 생성한 후 이를 기반으로 2010년 데이터에 대한 실험을 수행하였다. 먼저 Ψ 값이 1일 경우를 고려하였다. 이는 2009년 데이터의 확률분포 특성만을 반영한 돌연변이에 의한 자가 집합이다. 즉, 2009년 데이터 특성만을 반영했을 때 2010년 데이터의 특성을 분석하기 위함이다. 이를 통해 다음과 같은 결과를 생성하였다. 다음 결과는 자가 집합 각각에 대한 결과이다.

표 2. Ψ=1일 때 각 단위 자가 집합에 대한 진실 긍정비율
Table 2. True positive ratio base on self set when Ψ as 1

단위 자가 집합	진실 긍정(true positive)비율
수온	89%
염도	79%
DO	100%
pH	94%

3장에서 제시한 자가 집합을 활용한 해양환경 모니터링을 수행하기 위해, 2010년 데이터에 대해 100% 진실 긍정 비율을 가질 수 있는 Ψ를 결정하기 위한 실험을 수행하였다. 다양한 Ψ에 대한 결과는 다음과 같으며 질실 긍정이 100%인 각 단위 데이터의 Ψ의 값은 다음 표와 같다.

표 3. 진실 긍정이 100%일 때의 Ψ의 값
Table 3. Value of Ψ when true positive is 100%

단위 자가 집합	Ψ의 값
수온	3.2
염도	9.4
DO	0.7
pH	5.8

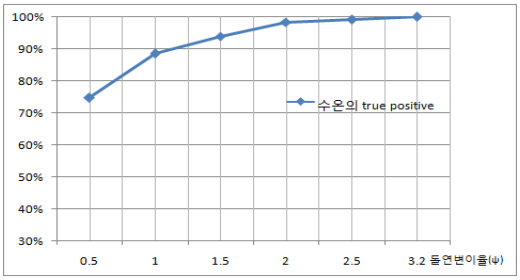

그림 10. Ψ에 따른 수온의 진실 긍정 비율
Fig. 10 True positive of water temperature base on Ψ

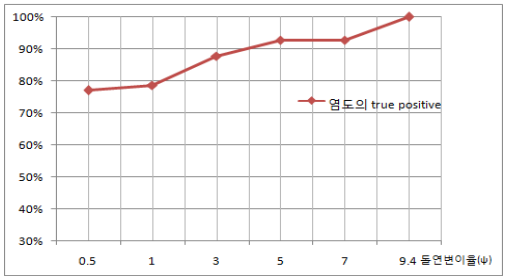

그림 11. Ψ에 따른 염도의 진실 긍정 비율
Fig. 11 True positive of salinity base on Ψ

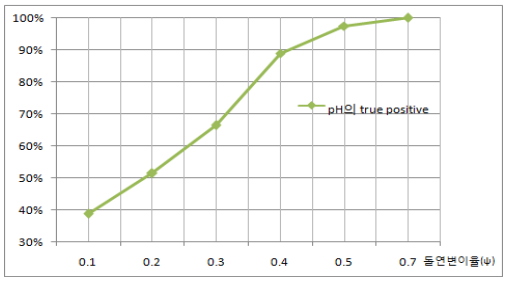

그림 12. Ψ에 따른 pH의 진실 긍정 비율
Fig. 12 True positive of pH Ψ

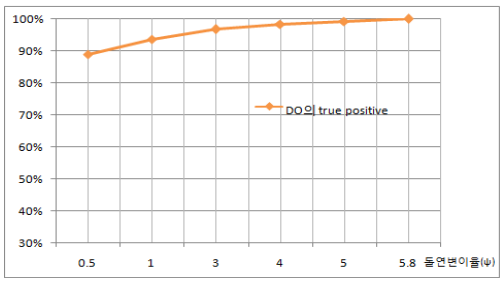

그림 13. Ψ에 따른 DO의 진실 긍정 비율
Fig. 13 True positive of DO to Ψ

제시한 결과를 통해, 2010년 데이터는 2009년에 비해 pH의 변화는 거의 없지만 염도, DO, 수온 순으로 변화가 많았음을 분석할 수 있다. 이는 2010년 데이터를 설명하기 위해 2009년 데이터의 특성을 얼마나 더 변화해야(즉, 돌연변이 해야)하는 지를 의미한다.

4.2 상태 정의 돌연변이의 자가 집합에 대한 해양환경 모니터링 실험

다음으로 해양 수질 데이터들에 대한 논리곱 관점에서 실험하였다. 이때, 해양 수질 인자들인 수온, 염도, DO, pH는 서로 다른 단위를 갖는 수치 정보이기 때문에 이를 정리하기 위해 Z-score로 변환하고 변환된 데이터에 대한 논리곱을 수행하고 이를 기반으로 상태 데이터 돌연변이를 수행하였다. 이때 획득한 자가 집합을 기반으로 2010년 데이터에 대한 모니터링을 수행하였다. 한편, 입력되는 데이터 또한 Z-score의 변환을 수행하였으며 이 과정에서 지속적으로 입력되는 데이터 관점에서 평균과 표준편차를 2009년 데이터를 활용하였다. 본 실험에서 2010년 데이터 전체를 알고 있지만 실제 환경에서 활용하기 위해 2009년의 평균과 표준편차를 활용하였다.

2009년 데이터에 대한 상태 데이터 돌연변이를 기반으로 생성한 자기 집합을 기반으로 2010년 데이터의 진실 긍정 비율을 평가하였다. 결과는 다음 그림과 같다. 앞선 실험과 달리 Ψ가 30일 때 진실 긍정 비율은 100%임을 알 수 있었다.

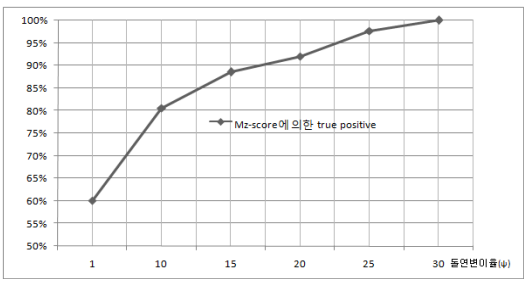

그림 14. Ψ에 따른 진실 긍정 비율
Fig. 14 True positive ratio to Ψ

각 해양 수질 요소들에 대한 논리곱을 기반으로 한 돌연변이 결과는 단위 데이터에 대한 돌연변이의 자가 집합에 비해 매우 높은 돌연변이를 요구하고 있다. 이는 수치 데이터의 논리곱에 의해 오차가 배가 됨으

로써 발생하는 것으로 추측한다.

Ⅵ. 결론 및 향후 연구

본 연구는 해양 USN 환경에서 수질환경의 온라인 정상·비정상 상태를 구분하기 위한 기법을 제안하였다. 이를 위해 자가 집합 정의가 필요하다. 수질환경의 정상 상태 집합을 정의하기 위해 인공면역시스템의 부정선택 메카니즘을 활용하였다. 센서로 수집된 대량의 해양 수질 환경 인자는 특정 상태를 의미하는 상수화 된 점 정보로 되어있기 때문에 정상·비정상에 대한 식별을 수행하기 위한 정보로서 부족하다. 이를 해결하기 위해 자가 집합 정의 과정에서 돌연변이를 적용하였다. 즉, 특정 상태정보를 범위로 모델링하기 위한 기법으로 단위 데이터 돌연변이를 통해 자가 집합을 생성하였다. 이를 위해 가우시안 함수를 기반으로 하였다.

가우시안 함수 및 돌연변이 상수 Ψ를 활용한 자가 집합 정의 방법으로, 본 연구에서는 두 가지 접근 방법을 제시하였다. 먼저 각 측정 인자에 대한 돌연변이를 통한 자가 집합 정의 방법이다. 다음으로 해양 수질 상태를 각 요소의 논리곱 관점에서 자가 집합을 정의 하였다. 직관적으로 후자가 타당하다고 추측될 수 있었으나 실험을 통해 제시한 바와 같이 전자의 방법이 최소화된 자가 집합을 필요로 하였다. 즉, 전자의 방법이 돌연변이 상수 Ψ가 낮았다. 이러한 정보처리는 현재 데이터를 기반으로 다음 데이터를 설명할 수 있는 Ψ 값에 따라 해당 수질 인자에 급격한 변화가 있었는지 또는 미미한 변화가 있었는지를 분석할 수 있었다. 따라서 지속적인 돌연변이 상수 Ψ의 관찰을 통해, 장기간 축적되는 해양수질환경 데이터의 Ψ 값을 분석함으로써 해양 수질 환경 변화추이를 관찰 할 수 있는 측정치로서 활용될 수 있다.

본 연구에서 제시한 자가 집합 정의를 통해 생성된 정보와 판단기법을 해양 수질 센서 로거단에 설치함으로써, 온라인으로 해양 수질 환경을 모니터링하고 비정상 인자를 식별 할 수 있다.

향후 연구과제로, 해양 UNS 환경에서 지속적인 Ψ 값의 변화량을 관찰하기 위한 더 많은 데이터 수집이 필요하다. 또한 수집된 데이터 분석을 하기 위한 기법으로 인공면역시스템의 클로널 선택 기법이나 위험이론 등을 고려한 수질 데이터의 특성을 분석할 계획이다.

참고 문헌

[1] 김경옥, 반경진, 허수연, 김응곤, “RFID/SUN 기반의 센싱 데이터 수집을 위한 시스템 설계 및 구현, 한국전자통신학회, 5권, 2호, pp. 221-226, 4, 2010.

[2] 표철식, 박상준, 김기일, 김동균, 이문규, 김관중, 김선진, 서정해, “훤히보이는 RFID/USN", 전자신문사, 2008.

[3] 이창복, “해양환경의 이해”, 서울대학교 출판문화원, 2011.

[4] 예성빈, 양수영, 정희택, “컴포넌트 기반 환경정보모니터링시스템 개발”, 한국전자통신학회, 7권, 1호, pp. 195-201, 2, 2012.

[5] 박혁, 이종수, “인공 면역 시스템을 이용한 설계 패턴의 인식 및 진화”, 대한기계학회, 25권, 11호, pp. 2421-2426, 2005.

[6] Leandro N. de Castro and Jonathan Timmis, “Artificial immune systems : a new computational intelligence approach”, Springer, pp. 1-29, 2002.

[7] 유경민, 양원혁, 이상열, 정혜련, 소원호, 김영천, “중앙집중형 망에서 인공체계 기반의 적응적 망 이상 상태 탐지 모델 설계”, 한국통신학회지, 34권, 3호, pp. 312-313, 3, 2009.

[8] 하기룡, 이도헌, “인공면역체계와 기계학습”,정보과학회지, 25권, 3호, pp. 76-82, 3. 2007.

[9] 심귀보, 황철민, “인공 면역 시스템과 분산 유전자 알고리즘에 기반한 자율 분산 로봇 시스템”, 한국통신학회지, 14권, 2호, pp. 164-170, 2004.

[10] 이상형, 김은태, 박민용, “Antibody Diversity 원리와 Antigen Presenting Cell을 구현한 새로운 인공 면역 시스템”, 전자공학회지, 41권, 4호, pp. 369-376, 7, 2004.

[11] 양병학, “새로운 최적화 기법소개 : 인공면역시스템”, IE Interfaces, 20권, 4호, pp. 458-468, 11, 2007.

[12] Hee-Taek Ceong, Hae-Ran Kim, "Estimation of Water Quality of Farms using Multivariate Statistical Analysis", KIMICS, Vol. 9, No. 4, pp. 475-481, 8, 2011.

[13] 유주형, 문정언, 민지은, 안유환 “정기여객선 현장관측 시스템과 Sea WiFS 자료를 이용한 서해 연안 해수환경 모니터링”, 원격감시한국저널, 23권, 4호, pp. 323-334, 2007.

저자 소개

정신출(Sin-Chul Jeoung)

2010년 7월~현재 전남대학교 대학원 디지털컨버전스 석사과정

※ 관심분야 : RFID/USN, 인공면역시스템

정희택(Hee-Taek Ceong)

1992년 2월 전남대학교 전산통계학과 학사

1995년 2월 전남대학교 전산통계학과 석사

1999년 8월 전남대학교 전산통계학과 박사

1999년 9월~현재 : 전남대학교 멀티미디어전공 교수

※ 관심분야 : RFID, 데이터마이닝, 분산처리시스템, 디지털 방송