실시간 임베디드 센서 네트워크 시스템에서 강건한 데이터, 이벤트 및 프라이버시 서비스 기술

(Robust Data, Event, and Privacy Services in Real-Time Embedded Sensor Network Systems)

정 강 수 [†] Krasimira Kapitanova [††] 손 상 혁 [†††] 박 석 [††††]
(Kangsoo Jung) (Krasimira Kapitanova) (Sang H. Son) (Seog Park)

요 약 실시간 임베디드 센서 네트워크 시스템에서의 이벤트 감지는 대부분 현실세계에서 수집된 센서 데이터들의 조합에 기반한다. 이에 최근에 이루어진 연구들에선 센서 데이터들을 수집, 집계하는 낮은 수준의 다양한 메커니즘들을 제안하였다. 그러나 실시간에서 연속적으로 발생하는 복잡한 이벤트들의 감지와 다양한 종류의 센서들로부터 입력되는 실시간 데이터의 처리를 위한 시스템에 대한 솔루션은 보다 많은 연구를 필요로 한다. 즉, 경량의 데이터 혼합이 가능하고 많은 컴퓨팅 자원을 필요로 하지 않는 실시간 이벤트 감지 기법이 필요하다. 이벤트 감지 프레임워크는 실시간 모니터링과 센서 데이터의 도착으로 일어나는 데이터 융합 메커니즘을 통하여 적시성과 임베디드 센서 네트워크의 자원 요구량을 감소시킬 수 있는 잠재력을 지니고 있다. 또한 임베디드 센서 네트워크 시스템이 신뢰성을 지닐 수 있도록 하기 위한 기반 기술인 프라이버시를 보장할 수 있는 익명화 기술을 설명한다.

키워드 : 임베디드 센서 네트워크, 이벤트 감지, 실시간 데이터 처리, 프라이버시

Abstract The majority of event detection in real-time embedded sensor network systems is based on data fusion that uses noisy sensor data collected from complicated real-world environments. Current research has produced several excellent low-level mechanisms to collect sensor data and perform aggregation. However, solutions that enable these systems to provide real-time data processing using readings from heterogeneous sensors and subsequently detect complex events of interest in real-time fashion need further research. We are developing real-time event detection approaches which allow light-weight data fusion and do not require significant computing resources. Underlying the event detection framework is a collection of real-time monitoring and fusion mechanisms that are invoked upon the arrival of sensor data. The combination of these mechanisms and the framework has the potential to significantly improve the timeliness and reduce the resource requirements of embedded sensor networks. In addition to that, we discuss about a privacy that is foundation technique for trusted embedded sensor network system and explain anonymization technique to ensure privacy.

Key words : Embedded Sensor Network, Event Detection, Real-Time Data Processing, Privacy

· 본 연구는 한국과학재단 세계수준의 연구중심대학(WCU) 육성사업(R33-2009-000-10110-0) 지원으로 수행되었음

† 학생회원 : 서강대학교 컴퓨터공학과
azure84@sogang.ac.kr
†† 비 회 원 : Univ. of Virginia Computer Science
krasi@virginia.edu
††† 비 회 원 : Univ. of Virginia Computer Science 교수
son@virginia.edu
†††† 종신회원 : 서강대학교 컴퓨터공학과 교수
spark@sogang.ac.kr

논문접수 : 2010년 8월 23일
심사완료 : 2010년 9월 30일

정보과학회논문지: 데이타베이스 제37권 제6호(2010.12)

1. 서 론

무선 센서의 계산 능력 향상과 지속적인 소형화에 따라 이를 사용하는 실시간 임베디드 시스템의 효용은 증가하고 있다. 임베디드 시스템들은 실세계를 모니터하고 적절한 대응과 제어를 제공하기 위하여 센서를 사용한다.

이러한 상호작용의 범위는 자원 제약적인 센서 기기부터 대규모의 모니터링 시스템에 이르기까지 매우 광범위하며, 인간이 실세계와 상호작용하고 제어하는 방식에 변화를 가져올 수 있는 잠재력을 지니고 있다. 실시간 임베디드 센서 네트워크 시스템들은 기간시설에 대한 모니터링, 의료 시스템과 스마트 헬스 케어, 감시용 설비, 그리고 환경에 대한 제어와 모니터링과 같은 다양한 분야에 사용된다.

실시간 임베디드 센서 네트워크에서는 대량의 가공되지 않은 데이터들이 끊임없이 수집된다. 따라서 가공되지 않은 데이터를 유용한 정보로 바꾸기 위해서는 여러 가지 데이터 및 이벤트 처리 기술이 필요하다. 예를 들어, 감시 시스템에서 사용되는 센서의 가공되지 않은 스트림 형태의 값들은 감시 지역 내의 모든 움직임과 관련된 의미 있는 이벤트들로 변환되어야만 한다. 따라서 잡음 섞인 실세계의 데이터를 처리하거나, 베이지안이나 뎀스터-쉐퍼 스킴과 같은 순수하게 확률적인 접근의 한계에 제약 받지 않는 새로운 추론 기술을 개발하여 데이터를 해석하고 유용한 정보를 구성하는 것은 주된 도전 과제가 된다.

또한 많은 수의 네트워크들이 대량의 실시간 센서 스트림 데이터를 제공할 때, 예측할 수 없는 미래에 대한 추론을 하는 것은 흥미로운 연구 과제이다.

신뢰성은 차세대 임베디드 센서 네트워크 기술의 주된 요구사항 중 하나이다. 프라이버시는 신뢰성의 주된 요소 중 하나이며, 이들은 프라이버시에 관련된 단원에서 논의될 것이다. 프라이버시와 더불어 신뢰성을 보장하기 위한 기반에는 스트림 데이터 처리에 대한 새로운 품질 관리(quality management)와 신뢰할 수 있는 데이터 혼합을 사용한 네트워크 내의 이벤트 탐지 기술이 필요하다. 기본적인 시스템 레벨에서의 역량 없이는 이후의 추론은 실패하거나, 손상된 데이터를 사용하여 연산을 수행하게 되므로 잘못된 결론으로 이어질 것이다. 만약 잘못된 결론이 액츄에이터를 움직이게 된다면 이는 심각한 안전 문제를 일으킬 수도 있다. 한 가지 접근법은 추론된 정보들이 확률적으로 신뢰할 수 있는 수준에서 수집되어 안전하게 액츄에이터의 동작을 보장하도록 하는 것이다. 또 하나의 접근 방법은 새로 생성된 정보를 사용하여 올바른 결정을 내리는 것이다. 그러나 결정을 내릴 때 거짓 양성(false positive)와 거짓 음성(false negative)의 수를 최소화하는 것과 안전성을 보장하는 것이 필요하다. 이를 보장하지 않는다면 시스템은 신뢰할 수 없다고 여겨질 것이다.

임베디드 센서 네트워크에서 대부분의 이벤트는 이진 형태가 아니다. 대신에 실세계의 복잡한 환경에 전개된 센서들의 데이터 혼합에 기반하고 있다. 가공되지 않은 센서 데이터들을 수집, 전송, 데이터 집계를 수행하기 위한 다양한 형태의 낮은 수준의 메커니즘과 프로토콜이 개발되었으나 센서 네트워크 시스템에서 실시간 데이터 프로세싱을 제공하기 위하여 다양한 종류의 센서로부터의 데이터 혼합과 연속적인 복잡한 이벤트를 탐지하는데 이용하는 시스템적인 솔루션에 대한 연구는 아직 충분하지 않다. 최근 센서 네트워크에서 신뢰할 수 있는 이벤트 탐지에 대한 연구가 이루어졌다[1]. 그러나 이러한 방식은 탐지된 이벤트가 현재의 위험을 나타내거나 쉽게 재현되기 힘든 상황(지진, 화산 폭발)의 경우에 사용되기 위해서는 적시적이고 예측할 수 없는 이벤트를 탐지하기 위한 센서 데이터의 적절한 변환이 필요하다.

임베디드 센서 네트워크를 위한 데이터, 이벤트, 그리고 프라이버시 서비스에서 반드시 만족되어야 할 몇 가지 요구사항들이 있다. 이벤트의 명세, 실시간 데이터 혼합, 실시간 스트림 데이터 관리, 그리고 프라이버시가 그것이다. 센서 기기들은 일반적으로 제한적인 자원을 가지고 있으므로 서비스는 경량화되어야 하며 높은 신뢰성 역시 제공되어야 한다. 이러한 모든 특성들은 임베디드 센서 네트워크 시스템을 위한 강건한 데이터, 이벤트, 그리고 프라이버시 서비스가 풀어야 할 난제들이다. 본 논문은 임베디드 센서 네트워크 시스템을 위한 강건한 실시간 데이터, 이벤트, 그리고 프라이버시 서비스를 제공하는 방법에 대해 이야기한다. 2장에서 복잡한 이벤트의 탐지를 위한 이벤트 표현 모델과 이벤트 변형 및 이벤트 서비스 프레임워크에 대해 소개하고 3장에서는 이벤트 처리의 대상이 되는 실시간 스트림 데이터의 특성과 품질관리에 대해 설명한다. 4장에서는 임베디드 센서 네트워크에서 수집되는 데이터에 대한 프라이버시 위협 문제와 이를 해결하기 위한 기법을 소개한다. 우리는 본 눈문에서 제안하는 기법이 이벤트 탐지의 적시성과 기반 메커니즘의 정확성에 있어 의미있는 개선을 이룸과 동시에 임베디드 센서 네트워크 어플리케이션의 프라이버시 요구사항을 만족시킬 것으로 기대한다.

2. 복잡한 이벤트 탐지

2.1 이벤트 표현 모델

Petri net은 동시적, 비동기적, 또한 비결정적인 특징들을 잘 묘사할 수 있는 모델로써 널리 사용되고 있다

[2]. Petri net은 기본적으로 위치(원형), 전이(직사각형, 혹은 바), 화살표, 그리고 토큰들(위치 안에 존재하는 점)로 구성된다. 화살표는 상태간의 변화와 어떤 토큰이 생성되고 소멸되는지를 표현한다. 위치는 어플리케이션이 취할 수 있는 상태를 나타내고, 전이는 다양한 종류의 액션들을 표현하기 위해 사용된다. 각각의 위치는 0개부터 여러 개까지의 토큰들을 포함할 수 있다. Petri net의 전이는 특정 토큰이 입력 위치에 존재할 때 촉발된다. 전이가 일어날 때, 토큰들은 소비되고 출력 위치에 토큰들이 삽입된다. Petri net의 마킹은 Petri net의 상태. 즉, 토큰의 분포를 나타낸다. 우리의 목적 중 하나는 Petri net에 기반하여 이벤트 감지를 위한 효율적인 명세 언어를 개발하는 것이다. 이러한 언어는 스트리밍 네트워크로부터의 데이터 분석에 기반하여 어플리케이션의 특징을 명세할 수 있어야 한다. 본 연구의 근간은 coMpact Event Description and Analysis Language(MEDAL)[3]이다. 정형화된 기법으로써 MEDAL은 Petri net에 기반하며, 복잡한 이벤트의 엄밀하고 명료한 명세를 가능하게 한다. MEDAL은 시간적 제어, 공간적 제약, 이종성, 그리고 확률적인 문제들과 같은 이벤트 감지 네트워크의 주된 측면들을 다룬다.

센서 네트워크 시스템 어플리케이션은MEDAL을 사용하여 7개의 튜플로 나타낸다: F = (P, T, A, λ, β, H, L), P는 위치들의 집합이고 T는 전이들의 집합, A는 화살표의 집합, λ는 화살표를 위한 확률/가중치 함수, β는 시간적 제약 함수, H는 위치를 위한 임계값 함수, L은 전이를 위한 공간적 제약 함수이다.

그림 1은 복잡한 이벤트 감지 어플리케이션의 MEDAL 모델을 보여준다. 온도 이벤트 1에 있는 토큰은 타임 t에 (x,y)에 위치한 센싱 범위 r을 지니는 센서에 의한 온도값 v의 감지를 나타낸다. 이벤트는 어플리케이션에 의해 감지되며, 이벤트 E는 온도(특정 장소의 높은 온도 감지), 압력(잠재적인 압력의 위험 수준의 감지), 마찰(접점에서의 빈번한 마찰의 발생)과 같이 명세된 간단한 이벤트들에 의해 기술된다.

각각의 간단한 이벤트들의 발생은 해당 위치에 있는 토큰에 의해 표현된다. 세 개의 토큰이 모두 존재할 때, 전이 T4가 촉발되고 어플리케이션은 이벤트 E의 감지를 보고한다. 복잡한 센서 네트워크 이벤트는 많은 경우 스트림 데이터의 처리를 필요로 한다. 이러한 이벤트들을 위해 토큰들은 스트림 데이터 처리에 대한 질의 결과를 사용하여 생성된다. 이것에 대하여는 제3장에서 다룰 것이다.

복잡한 이벤트는 언제, 그리고 어디에서 일어나는지의 함수이다. MEDAL내에서, 시간적이고 공간적인 의미를 고려한 이벤트 감지 시스템의 동작은 다음과 같다.

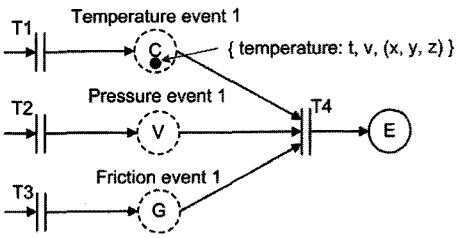

그림 1 MEDAL model of an event detection system

2.1.1 Temporal logic

시간적 로직은 시간적 제약 함수 β를 참조한다. 이는 MEDAL Petri net 안에서, 시간적 개념인 '언제'와 '얼마나 오래'를 명세하는 데 도움을 준다. β는 오직 명세된 시간 간격동안만 전이가 일어나는 것을 보장한다. MEDAL내에서 β의 도입은 현실적인 중요성을 지니는데, 어떤 이벤트들은 오직 특정 시간간격 동안에만 일어날 수 있기 때문이다. 예를 들어 클러스터링은 오직 상대적으로 짧은 시간동안 집중적으로 통신이 발생할 때에만 발생한다. β는 '전이는 오직 입력 토큰이 사전에 정의된 시간 간격 안에서 생성될 때에만 발생한다'와 같은 조건을 명세하는데에도 도움이 된다. 예를 들어 그림 1에서 T4로 진입하는 토큰들간의 발생 시간이 30초 이상이라면, 이것은 이벤트들의 서로 상관이 없거나, 서로 다른 그룹에 속해있을 확률이 높음을 의미한다. 이와 같은 경우에 네트워크는 필요한 입력 토큰이 존재해도, 이벤트의 발생을 보고해서는 안 된다.

2.1.2 Spatial logic

어플리케이션의 지리적 의미는 공간적 함수 L에 의해 수행된다. 전이 T를 위한 제약 함수로써 L은, T로 향하는 화살표들이 공간적인 장소 조건을 만족시킬 때 토큰이 운반되는 것을 보장한다. 만약 L(T)=R이라면, 상위-레벨 이벤트의 효율적인 범위의 반지름 T가 반지름 R보다 같거나 작아야만 인정된다. 달리 말하면 하나의 특정한 이벤트를 발생시키기 위한 데이터를 고려하기 위해서는 반지름 R은 모든 토큰들의 위치를 포함해야 한다. 예를 들어 그림 1에서, 만약 전이 T4의 input 위치의 토큰들이 서로 R보다 더 먼 거리에 있다면, 이벤트들끼리 별다른 상관이 없다는 것을 의미한다. L은 이러한 상황을 감지하고 거짓-양성(false-positive)의 수를 줄이는데 도움을 준다.

2.2 퍼지 로직 기반의 확률적 이벤트 탐지 모델

임베디드 센서 네트워크에서의 이벤트 기술에 대한 대부분의 이전 연구들은 정확한(crisp) 값을 이벤트의 특징을 명세하기 위한 파라미터로 사용하였다. 예를 들어 온도가 5°C 이하로 떨어지는 것을 탐지하기를 원한다고 가정해보자. 문제는 센서가 늘 정확한 값을 읽지

않는다는데 있다. 또다른 문제는 서로 근접한 위치에 있는 센서라 해도 다른 값이 입력될 수 있다는 것이다. 만약 온도가 5°C 이상으로 올라갈 때 냉방이 작동하기를 원하는 시나리오를 생각해보자. 두 센서 A, B가 방안의 온도를 측정하고 그들이 보고한 값의 평균을 이용하여 액션의 수행 여부가 결정된다. 어떤 시점에서 센서 A는 5.1°C를, 센서 B는 4.8°C를 보고하며 이때의 평균은 4.95°C이다. 이는 사전에 정의된 임계값 이하이므로 냉방은 작동되지 않는다. 그러나 만약 센서 B의 측정이 정확하지 않아서 실제 온도보다 낮은 값을 측정했다면 우리는 잘못된 결론을 내리게 된 것이다. 이 상황은 두 개 이상의 센서 측정이 연관될 때 더욱 복잡해진다. 이는 정확한 이벤트 임계값을 정확한 값을 사용하여 우리가 믿을 수 있도록 결정하는 것이 어려운 일이고 최상의 접근이 아님을 알 수 있게 한다. 반면에 퍼지 로직은 센서 네트워크에서의 이벤트를 묘사하는데 보다 적합한 다음의 여러 가지 특징들을 지니고 있다:

- 부정확하고 신뢰할 수 없는 센서 입력에서도 잘 동작할 수 있다.
- crisp 로직보다 인간의 생각에 가깝다. 예를 들어, 퍼지 로직은 화재에 대한 이벤트 묘사를 '높은 온도와 연기의 존재'로 생각한다. 이는 이벤트의 특징을 '55도씨 이상의 온도와 연기에 의한 어두움이 15%이상'인 것보다 인간의 생각에 더 유사하다.
- 확률에 기반한 다른 분류 알고리즘과 비교해서 퍼지 로직은 보다 직관적이고 사용하기 쉽다.

일반적인 퍼지 로직 시스템(FLS)의 구조는 그림 2와 같다. 첫째, 퍼지파이어(fuzzifier)는 입력되는 crisp의 변수 x ε X(이 때 X는 가능한 입력 변수들의 집합이다)를 대응하는 멤버쉽 함수를 사용하여 퍼지 언어적 변수로 변환한다. 언어적 변수는 "숫자가 아닌 자연언어나 인공 언어의 단어나 문장을 값으로 하는 변수"이다[4]. 입력 변수들은 계산되어진 멤버쉽의 정도에 따라 하나 나 둘 이상의 퍼지 셋들과 연관된다. 예를 들어 온도값은 '차가움'이나 '미온'으로 분류될 수 있다. 둘째, 퍼지화된 값들은 규칙(rule)이라 불리는 if-then 언어 문장에 의해 처리된다. 이는 전문가가 제공한 도메인 지식으로부터 유래한다. 이 규칙은 다음의 형태를 지닌다:

IF premise, THEN consequent

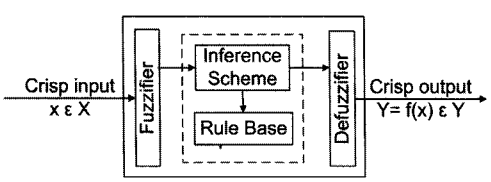

그림 2 The structure of a fuzzy logic system

premise는 논리 함수(e.g.AND, OR, NOT)에 의해 연결되는 퍼지 입력 변수들의 조합으로 이루어진다. 그리고 consequent는 퍼지 출력 변수이다. 이 추론 스킴은 입력 퍼지 셋과 출력 퍼지 셋을 대응시킨다. 마지막으로 디퍼지파이어(defuzzifier)는 출력퍼지 셋들을 사용하여 crisp 출력을 연산한다. crisp출력 값은 시스템이 취해야 할 제어동작을 결정한다.

앞에서 언급했던 것과 같이, 센서에서 읽어들인 값은 일반적으로 신뢰할 수 없고 부정확하다고 간주되어진다. 그러므로 관찰 영역에서의 이벤트의 발생에 대한 신뢰도를 향상시키기 위해서, 우리는 자주, 여러 센서로부터 일정 주기에 걸쳐 여러번 값을 읽어들이는 것이 필요하다. 이를 위해 우리는 공간적이고 시간적인 의미를 지니는 이벤트 감지 프로세스를 제안했다. 이러한 방식은 규칙-기반 안에 시간적이고 공간적인 변수를 포함시키는 것이 감지의 정확성을 크게 향상시킬 수 있을 것이라는 논리에 근거한다. 더 나아가서 이러한 방식은 보다 복잡한 이벤트를 감지하고 기술하는 것을 가능하게 한다. 우리가 알고 있는 한, 이벤트 감지의 정확성을 위해 시간적이고 공간적인 의미의 효과를 고려하여 퍼지 로직에 적용한 연구는 이제까지 깊이 있게 다루어지지 않고 있다.

2.2.1 Spatial semantics

공간적 제약을 포함시키는 것이 이벤트 감지의 정확도에 어떻게 영향을 미치는지 이해하는 것이 중요하다. 공간적 제약 변수는 공간적인 요구사항을 탐지 어플리케이션이 표현하는 것을 가능하게 한다. 다음과 같은 요구사항이 존재할 수 있다. "만약 두 센서 노드로부터 읽어들인 값이 이벤트 X의 출현을 나타내는 경우 우리는 두 센서 노드가 서로 가까이 위치했을 때만 이 이벤트가 일어났다고 믿을 수 있다." 공간적 의미를 포함했을 때의 단점은 이벤트의 공간적 제약이 너무 엄격해서 때론 이벤트가 감지되지 않을 수 있다는 것이다. 그러나 퍼지 로직이 이런 문제점을 완화하는데 도움이 될 것으로 기대된다.

2.2.2 Temporal semantics

거짓 경보를 감소시키기 위해서는 감지되는 이벤트들의 시간적 특성을 측정하는 것이 필요하다. 한 가지 접근법은 시간적 제약처럼 작동하는 규칙 기반의 언어적 변수를 포함시키는 것이다. 시간적 의미를 더하는 것은 센서 통신의 특성상 임베디드 센서 네트워크에서 특히 중요하다. 무선 센서 네트워크에서는 네트워크 혼잡이나 라우팅 문제로 인해 메세지의 지연이 발생할 수 있다. 따라서 신뢰할 수 있는 규칙 기반의 이벤트 감지는 센서에서 읽어들인 값의 생성 시간을 고려해서 이루어져야 한다.

2.2.3 Decreasing the rule-base

퍼지 로직을 사용하는 것의 단점은 규칙을 저장하는

것이 상당히 많은 양의 메모리를 필요로 한다는 것이다. 룰의 숫자는 변수의 숫자에 지수적으로 증가한다. n개의 변수들이 m개의 값을 지닌다면 규칙-기반에서 규칙 의 갯수는 m^n이다. 예를 들어, 만약 4개의 언어적 변수가 각각 5개의 값들을 취한다면 규칙-기반은 625개의 규칙을 갖게된다. 공간적이고 시간적인 언어적 변수들을 더 한다면 이 숫자는 더욱 증가하게 된다. 센서는 제한된 메모리를 가지므로 전체 규칙을 모든 노드에 저장하는 것은 가용 자원을 낭비하는 것이 될 수 있다. 게다가, 지속적으로 많은 규칙을 순회하는 것은 감지 프로세스를 느리게 만든다. 또한 많은 규칙에 의해 일어나는 증가된 계산은 전력의 소비를 높인다. 이 문제를 해결하기 위해 우리는 규칙의 크기를 줄이기 위한 규칙-기반 감소 기법의 집합을 설계했다[5]. 이 기법의 중요한 특성은 정확도와 이벤트 감지의 신속성에 부정적인 영향을 끼치지 않는다는 것이다.

2.2.4 실험 결과

우리는 National Institute of Standards and Technology(NIST) 사이트에 공개된 실제 화재 데이터를 사용해 시뮬레이션을 수행했다. 실험의 목적은 퍼지로직을 사용한 것이 어떻게 이벤트 감지의 정확도에 영향을 미치는지 평가하는 것이다. 실험은 crisp와 퍼지 값, 두 종류에 대해 이루어졌다. 실험에 사용한 임계값은 상용 연기와 열 감지기를 사용하였다[6,7].

그림 3은 crisp값에 대한 실험 결과의 하나이다. 동기화 시스템의 시작은 점화 시간으로 표현된다. 우리가 그림에서 볼 수 있듯이, crisp값을 사용하는 것은 많은 횟수의 거짓 화재 감지를 보인다. 화재가 일어나기 전의 기간동안 약 50번의 잘못된 화재 감지를 기록하였으며 이는 전체 입력의 1.3%에 해당한다. 이러한 거짓-양성 결과는 전체 이벤트 감지 시스템의 효율성과 신뢰도에 심각한 영향을 미친다. 이 결과는 퍼지 로직이 센서 네트워크에서의 이벤트 감지에 사용되는 것이 보다 적합함을 보여준다. 그러나 보다 복잡한 어플리케이션에서도 이러한 적합성을 유지하지 못 한다면 추가적인 연구가 필요하다. 또한 어플리케이션의 퍼지와 crisp를 사용했을 때의 자원 요구량을 비교하는 것도 중요하다. 우리는 같은 시나리오를 사용하여 규칙-기반의 감소 기법의 효율성을 평가하였다. 규칙-기반은 81개의 규칙으로 초기화하였다. 우리의 감소 기법중 두 개를 적용한다면 이벤트 감지의 정확성을 손상시키지 않으면서도 규칙의 크기를 70%가량 감소시키는 것이 가능해진다[5].

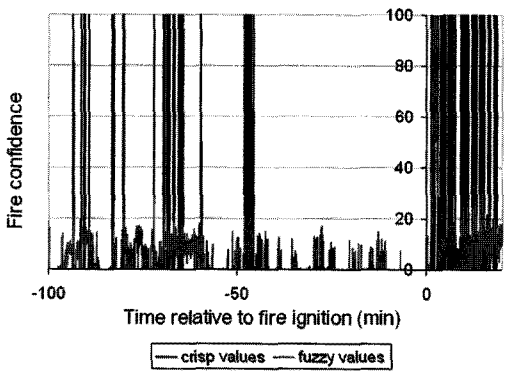

그림 3 Fire simulation: burning mattress

2.3 이벤트 변형

센서 노드에서 실행되는 정형 모델로의 전환은 이벤트 서비스에서 필요로 하는 다음 단계이다. 명세된 이벤트의 처리가 DNA의 복제 과정과 유사하므로 MEDAL 모델로부터의 이벤트 인지 코드 생성과 센서 노드에의 저장을 이벤트-DNA라 부른다. 개개의 이벤트-DNA는 MEDAL 모델의 인코드된 표현이며 MEDAL 모델과 같이 간단하거나 복잡한 이벤트의 기술을 표현한다. 센서 노드는 메모리에 서로 다른 이벤트-DNA를 읽어들이고 그에 맞는 동작을 저장하는 이벤트 감지 미들웨어를 지닌다.

2.4 이벤트 서비스 프레임워크

그림 4는 우리가 제안하는 이벤트 서비스 구조이다. 그림에서 MEDAL IDE(Integrated Development Environment)는 오프라인 패키지로 PC에 위치하며 명세된 이벤트의 의미를 인코드하고 Base 노드로 내보낸다. Base 노드는 해당하는 센서 모트에 전체 이벤트-DNA들을 전개하는 전개 모듈을 설치한다. 예를 들어 만약 이벤트-DNA가 모트 수준의 이벤트를 표현한다면 이는 모든 모트들에 전개되며, 만약 이벤트 DNA가 그룹 레벨의 이벤트로 표현된다면 그룹의 리더들만이 복사본을

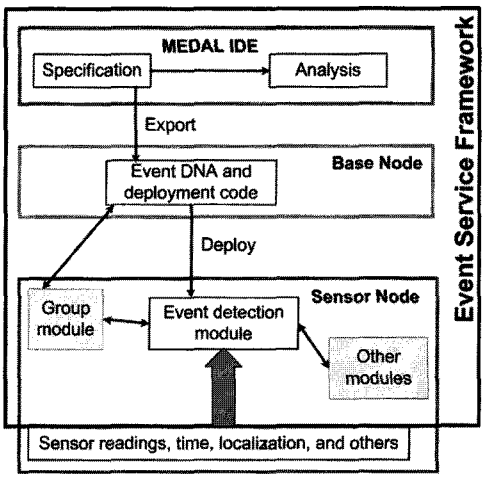

그림 4 Event service framework

갖는다. 동적인 그룹 리더 스킴에서 전개 모듈은 무엇을 전개할지 결정하기 위해 그룹 모듈과 상호작용하는 것이 필요하다. 그 후 모트들마다 이벤트-DNA로부터 내재된 코드를 생성하여 프로그램 메모리에 저장한다. 감지 모듈은 읽어들인 센서 값이나 획득 시간, 위치 정보, 그 외에 프로그램 섹션 내의 다른 모듈로부터 얻어지는 정보와 같은 낮은 수준의 호출을 통해 그 목적을 이룬다.

프레임워크는 다음의 모듈들로 이루어진다.

- MEDAL environment, 명세 모듈과 분석 모듈을 포함한다. 이벤트 DNA안의 이벤트 명세를 인코드하고 이를 Base 노드로 내보낸다.
- Base node는 인코드된 이벤트-DNA와 이벤트-DNA를 필드의 노드로 전송하기 위한 전개 모듈을 포함한다. 전개 모듈은 통신을 제공하고 이 때의 메시지는 이벤트-DNA이다.
- Sensor node 레벨에서, 이벤트 감지 모듈은 프로그램 메모리 안에 있다. 이벤트 감지 모듈은 상위 레벨의 이벤트의 협업 감지를 위해 토큰 벡터를 사용하여 다른 노드와 통신한다.

일단 설계자가 어플리케이션을 기술하기 위해 MEDAL 모델을 사용하면, 다음 단계는 노드에서 실행하기 위한 코드를 작성하는 것이다. 이 단계의 약점은 정형 모델을 직접 변환하는 것이 모델로부터 생성된 코드와 불일치를 일으켜 버그처럼 보일 수 있다는 것이다. 이벤트 기술을 위해 사용되는 다른 접근에 비해 MEDAL 모델이 지니는 장점은 정형화된 구조와 명료함에 있고, 이는 즉각적이면서 자동적으로 센서 노드에서 실행되는 코드로 변환이 가능하다. 우리는 MEDAL 모델에 기반하여 이벤트-DNA 코드를 자동으로 생성하는 툴을 개발하였다. 현재 우리는 정형화된 어플리케이션 모델[1]을 사용하여 TinyOS코드를 생성하고 있다. 이러한 방식은 nesC를 사용하는 것보다 더 적응성이 좋다. 이 방식은 TinyOS 코드를 작성하는 노력을 크게 줄여주며 코드의 정확성을 향상시킨다.

3. 실시간 스트림 데이터

적시의 복잡한 이벤트 감지를 위하여 실시간 임베디드 센서 네트워크 시스템은 여러 소스로부터 연속적으로 무제한 데이터 스트림에 대한 작업을 수행해야 한다. 서로 다른 형태의 센서로부터 들어오는 스트리밍 데이터는 서로 결합되어 처리되어야 하며 이는 이미 시스템에 저장되어 있는 데이터와도 마찬가지이다. 스트림 데이터는 연속적이고, 순차적이며, 잠재적으로 무한한 데이터 스트림의 형태를 취하는데 이는 유한하고, 정적인 저장된 데이터들과는 반대이다. 임베디드 센서 네트워크 어플리케이션에서는 데이터 스트림의 형태로 존재하는 현재의 위치나 움직임, 다른 하부조직에 의해 감지된 단순한 이벤트, 그리고 데이터 스트림의 동적이고 일시적인 특징 등의 동적으로 변화하는 데이터가 상당히 많은 분량을 차지하고 있다. 스트림 데이터의 분석은 실시간 데이터 관리의 중요한 문제인데 이는 방대한(잠재적으로는 무한한)양, 예측할 수 없는 패턴의 변화, 그리고 단기간에 동적으로 변하는 순서 등의 데이터 스트림의 독특한 특성 때문이다. 데이터 스트림의 방대한 양과 어플리케이션의 시간적 제약에 의해, 스트림 데이터 전체를 저장하는 것은 불가능하며, 전체 스트림 히스토리에 대한 질의를 하는 것 역시 적절치 않다고 여겨진다. 일반적으로 질의는 데이터의 윈도우에 대해 실행된다. 윈도우 상의 데이터 스트림은 현재의 질의를 위해 고려되어지는 데이터 스트림의 부분이다. 많은 스트림 데이터는 가공되지 않은 센서 데이터의 형태로 존재한다. 이는 집계를 수행하거나 가공되지 않은 데이터로부터 적절한 추상화 레벨과 차원의 조합을 통해 흥미로운 패턴이나 특별한 데이터를 추출하기 위해 필요하다

3.1 실시간 스트림 데이터 서비스 품질 관리

임베디드 센서 네트워크 안에서의 실시간 스트림 데이터를 처리하기 위해서, 시스템은 오랫동안 지속되는 질의를 지원할 수 있어야 한다. 질의가 시스템에 도착하면, 질의는 등록되고 주기적으로 그 인스턴스가 실행된다. 모든 질의들은 시스템에 예비 등록되고, 실행되기 전에 질의 계획(연산자, 큐, 시놉시스를 포함하는)으로 변환된다. 질의 계획 안에서 큐는 입력 데이터 스트림과 연산자 간의 중간 결과를 모델링한다. 시놉시스는 명세된 연산자와 시스템의 추후 평가를 위한 상태를 저장하는 것과 관련된다. 예를 들어 조인 연산자는 몇 가지 아이템을 저장하기 위한 입력값들을 위해 시놉시스와 관련된다. 이 아이템들은 필요할 때 연산자에 의해 탐색된다.

아래의 데이터 스트림과 관련된 질의들을 고려해보자. 질의 결과는 비정상적으로 고속으로 이동하는 트럭을 탐지하기 위해 사용된다.

Stream: Speed (int lane, float value, char[8] type);
Relation: Lanes (int ID);
Query: SELECT avg (Speed.value) FROM Speed [range 1 minute], Lanes WHERE Speed.lane = Lanes.ID AND Speed.type = Truck;
Period 10 seconds
Deadline 5 seconds

위의 데이터 스트림에 대한 질의 연산은 속도 센서와 마지막 1분에서 특정 차선에 있는 평균 트럭의 속도를 계산한 평균에 의해 생성된다. 질의는 매 10초마다 수행될 것을 필요로 하며 데드라인은 주기적인 질의 인스턴스가 수행된 후 5초 이내이다. 생성된 질의 계획은 그림 5

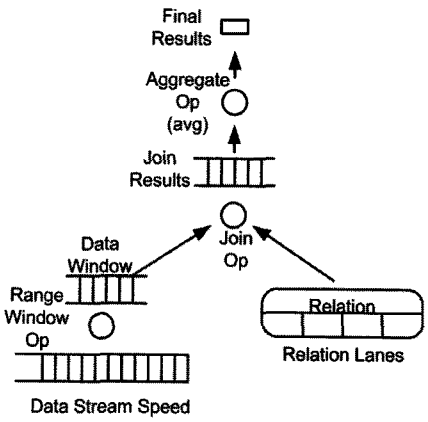

그림 5 Query plan

와 같다. 이 질의 계획은 3개의 질의 연산자와(범위 윈도우 연산자, 조인 연산자, 집계 연산자)두개의 버퍼(범위 윈도우 결과를 저장하기 위한 버퍼와 조인 연산자의 결과를 저장하기 위한 버퍼)로 이루어진다.

실시간 제약이 있는 데이터 스트림 관리에서의 주된 문제는 임베디드 센서 네트워크와 같이 데이터 스트림 자체의 예측불가능함이다. 시스템은 도착율과 입력 데이터 스트림의 내용의 변화에 따라 과부하가 걸릴 수 있다. 시스템은 이러한 업무량의 변동을 다룰 수 있어야 한다. 만약 그렇지 않으면 질의 중 몇몇은 데드라인을 넘길 수 있다. 서비스 품질(QoS) 관리는 두 단계로 수행된다. inter-query QoS 관리는 질의 결과가 비슷한 품질을 나타내도록 하기 위해서 시스템이 과부하가 걸리는 경우에 서로 다른 질의에 자원을 할당한다. intra-query 관리는 질의의 품질이 최대화가 되도록 질의 계획내의 서로 다른 연산자에 가용 자원을 배치한다. inter-query QoS 관리를 위해서 시스템은 서로 다른 입력 데이터 크기와 대응하는 쿼리 수행 시간을 측정할 필요가 있다. intra-query QoS 관리를 위해서 시스템은 현재의 데이터와 대응하는 연산자들의 선택성과 그들의 수행 시간의 측정값을 아는 것이 필요하다. 우리의 연구 목적은 질의 결과의 적시성을 보장하는 것이므로, 질의 수행 시간 평가는 서비스 품질 관리의 주된 요소이다. 질의 수행 시간을 측정하기 위해서는 각 질의마다의 입력 데이터 스트림의 양, 연산자의 선택성, 그리고 각 연산자마다의 데이터 튜플당 수행 시간이라는 세 가지 인자를 필요로 한다. 또한 우리는 QoS 관리 루틴이 실행될 때, 실행될 준비가 완료된 질의를 고려해야 한다. 그러므로 이러한 질의들을 위한 입력 데이터의 양은 입력 데이터 스트림의 세그먼트가 이미 시스템에 존재하는 시점으로부터 알려진다.

아직 완전한 입력값을 가지지 않은 질의의 수행 시간을 측정하거나 서비스 품질 관리 프로세스에서 이 측정값을 사용하는 것은 현재 실행 직전의 질의를 미래의 질의와 부분적으로 중복시킬 수 있으므로 유용하다. 이는 데이터 스트림의 양과 내용을 측정하고 모니터하는 효과적인 알고리즘을 설계하는 것을 포함하는 중요한 문제이다.

4. 임베디드 센서 네트워크에서 프라이버시 서비스

임베디드 센서 네트워크가 확산됨에 따라 방대한 양의 가공되지 않은 정보들이 실시간으로 수집되면서, 개인의 프라이버시에 대한 우려가 제기되었다. 예를 들어, 거리에 설치된 비디오 시스템이나 스마트 헬스 케어를 목적으로 설치된 센서 장비에 의해 수집되는 개인의 현재 위치나 건강 상태 등의 데이터는 민감한 사적인 정보와 직접적으로 연관된다. 만약 수집된 데이터를 서비스 제공자가 악용하거나 실수로 유출될 경우 심각한 프라이버시의 침해가 발생할 수 있다. 이러한 프라이버시에 대한 우려는 임베디드 센서 네트워크를 실생활에 적용함에 있어 중요한 문제이다. 사용자들이 시스템의 안전성과 보안성을 신뢰하지 않는다면 그 시스템은 사용되지 않을 것이다. 이 장에서 우리는 임베디드 센서 네트워크에서의 프라이버시 보호 기법과 구조에 대해 설명하기로 한다.

4.1 익명화 기법

정적인 데이터를 대상으로 개인의 프라이버시를 보장하기 위한 연구들이 이루어져 왔다. 가장 일반적으로 사용되는 것은 K-anonymity[8]와 같은 익명화 기법으로써 사용자의 정보가 노출되는 것을 피하기 위하여 같은 식별 정보를 지니는 사용자를 k-1명 이상 유지함으로써 사용자가 식별될 수 있는 확률을 1/k로 낮추는 것이다. 또한 인터넷과 같은 유선 네트워크에서 송신자의 익명성을 보장하기 위한 기법들도 연구되었다[9]. 그러나 기존의 프라이버시 보호 기법들은 에너지 제한적이고 시간적, 공간적 제약 사항이 따르는 무선 센서 네트워크 환경에 적용하기에는 무리가 있다. 따라서 무선 센서 네트워크의 특성을 고려한 경량화되고 자원 소비가 적은 새로운 프라이버시 보호 기법에 대한 연구의 필요성이 대두되었다.

4.2 익명화 구조

익명화를 수행하는 구조는 크게 2가지로 분류할 수 있다. 하나는 중앙 집중화된 신뢰할 수 있는 익명화 서버를 이용하여 수집된 정보를 익명화하여 개인의 식별 정보가 노출되지 않도록 하는 것이다.

그림 6은 중앙 집중형 익명화 구조를 보여준다. 이 방법은 신뢰할 수 있는 서버를 사용하여 익명화를 수행하

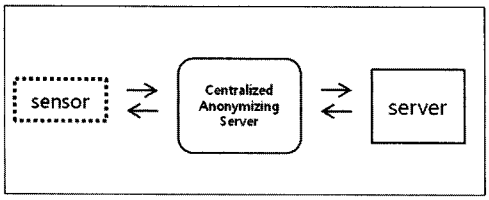

그림 6 중앙 집중형 익명화 구조

므로 센서 노드에 걸리는 부하가 없다는 장점을 지니지만 병목 현상이 발생할 수 있으며 하나의 장소에서 모든 익명화 과정을 수행하므로 서버가 공격당할 경우 피해가 크다. 또 다른 방법은 ad-hoc 네트워크에 기반한 것이다. 이 방법은 센서 디바이스를 사용하여 서버로 정보를 보내기 전 주변에 있는 센서들과 네트워크를 구성하여 익명화 과정을 수행하는 것이다.

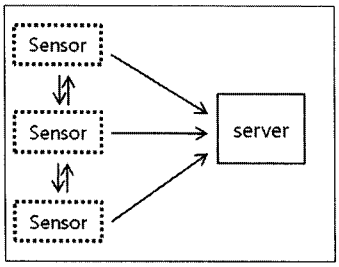

그림 7 ad-hoc 기반의 익명화 구조

그림 7은 ad-hoc네트워크를 이용한 익명화 구조를 보여준다. 이 방법은 확장성과 별도의 신뢰할 수 있는 익명화 서버를 두지 않아도 된다는 장점을 가지고 있으나 자원 제약적인 환경에서 센서에 과도한 부하가 걸릴 수 있다는 단점이 있다.

중앙 집중형과 분산형 익명화 구조는 장단점이 존재하므로 센서 네트워크가 사용되는 환경에 맞는 구조를 선택하는 것이 중요하다. 우리는 중앙 집중형 구조와 분산형 익명화 구조를 결합한 프라이버시 보호 모델을 제안한다.

4.3 중앙 집중과 분산형 결합 익명화 모델

그림 8은 제안 기법의 익명화 구조를 나타낸다. 중앙 집중형 익명화 구조의 단점은 서버를 신뢰해야 한다는 제약과 집중된 연산으로 인한 병목 현상이고 ad-hoc 네트워크를 이용한 익명화 구조의 단점은 자원 제약적인 센서 노드에 과도한 부하가 걸릴 수 있다는 점이었다. 본 제안 모델은 데이터 연산을 수행하는 서버와 센서 노드 사이에 서드-파티 익명화 관리자를 둠으로써 이 문제를 해결했다. 서드-파티 익명화 관리자는 각각의 센서 노드에 대한 메타 데이터를 보관하고 이를 바탕으로 센서로부터 수집되는 데이터에 대한 효율적인 익명화를

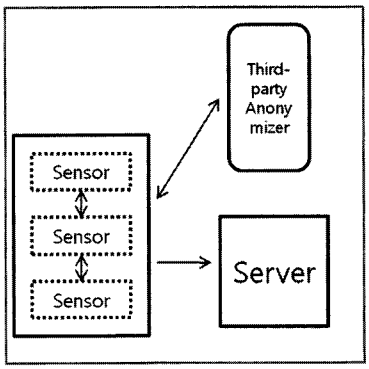

그림 8 제안 기법의 익명화 구조

유도한다. 센서들은 익명화 관리자와의 통신을 통해 익명화를 수행하고 익명화 된 데이터를 서버로 전송한다. 익명화 관리자는 익명화에 필요한 데이터 외의 정보를 요구하지 않으므로 신뢰할 필요가 없으며 센서는 익명화 관리자가 제공하는 메타 데이터에 기반하여 익명화 과정을 수행하므로 센서간 통신에 드는 비용을 줄일 수 있다. 서버는 센서로부터 익명화된 데이터를 받아 데이터 처리를 수행하므로 병목현상 및 집중화된 서버에 대한 공격의 위험을 경감시킬 수 있다. 제안 구조는 자원 제약적이고 동적으로 변화하는 임베디드 센서 네트워크에서 필요한 프라이버시 요구사항을 만족시킬 수 있을 것으로 예상하며 실시간 데이터 스트림 기법과 이벤트 감지의 기반이 되는 신뢰성 있는 시스템을 구축하는데 중요한 역할을 수행한다.

5. 요 약

이 논문에서 제안된 방식들은 감시와 감지, 그리고 정교한 이벤트에 대한 반응을 위한 센서 네트워크에서에서의 강건한 실시간 데이터와 이벤트 서비스를 개발하기 위한 기반을 제공한다. 이 논문은 임베디드 센서 네트워크에서 경량화되고 이벤트 감지의 적시성을 제공하는 몇 가지 기법을 결합하여 이벤트 서비스 프레임워크를 제안하였다. 추가적으로 이 센서 시스템의 성능을 향상시키기 위하여 서비스 품질 관리 기법을 개발하였으며 신뢰성 있는 시스템을 위한 프라이버시 보호 모델을 제안하였다. 우리는 이벤트 서비스 프레임워크와 서비스 품질 관리 메커니즘의 결합은 임베디드 센서 네트워크 시스템에서 이벤트 감지를 위한 자원 요구를 감소시키고 적시성을 향상시키는데 중요한 역할을 할 것으로 기대되며, 개인 정보 노출 등의 개인 프라이버시에 대한 안전성을 보장하는 것이 차세대 임베디드 센서 네트워크의 발전을 위해 중요할 것으로 생각한다.

참 고 문 헌

[1] M. Keally, G. Zhou, G. Xing, "Watchdog: Confident Event Detection in Heterogeneous Sensor Networks," *RTAS2010.*

[2] C. Girault, R. Valk, "Petri Nets for System Engineering: A Guide to Modeling, Verification, and Applications," Springer-Verlag New York, Inc. (2001).

[3] K. Kapitanova, S. H. Son, "MEDAL: A coMpact Event Description and Analysis Language for Wireless Sensor Networks," *International Conferenceon Networked Sensing Systems*, 2009.

[4] L. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Transactions on Systems, Man, and Cybernetics*, 1973.

[5] Krasimira Kapitanova, Sang H. Son, and Kyoung-Don Kang, "Event Detection in Wireless Sensor Networks-Can Fuzzy Values Be Accurate?," Second International Conference on Ad Hoc Networks, 2010.

[6] J. Geiman and D. Gottuk, "Alarm Thresholds for Smoke Detector Modeling," *Fire Safety Science - Proceedings of the 7th International Symposium* 2002.

[7] WS4916 Series Wireless Smoke Detector. http://www.alarmsuperstore.com/dsc/ws4916installation.pdf.

[8] L. Sweeney, "k-anonymity: a model for protecting privacy," International Journal on Uncertainty, Fuzziness, and Knowlege-Based Systems, 2002.

[9] Michael K. Reiter, Aviel D. Rubin, "Anonymity for Web Transactions," ACM Transaction on information and system security, 1998.

정 강 수

2007년 서강대학교 컴퓨터학과(공학사)
2009년 서강대학교 컴퓨터학과(공학석사)
2007년 9월~현재 서강대학교 컴퓨터공학과 박사과정 재학중. 관심분야는 프라이버시, DaaS, 접근제어임

Krasimira Kapitanova

Krasimira Kapitanova is currently working towards a Ph.D. in Computer Science at the University of Virginia. She received her B.S. degree in Computer Science and Technologies from Technical University Sofia, Bulgaria, and an M.C.S. degree from the University of Virginia. Her research interests include formal event description in sensor networks, QoS management, and information management and security.

손 상 혁

Sang Hyuk Son is a Professor at the Department of Computer Science of University of Virginia. He received the B.S. degree in electronics engineering from Seoul National University, M.S. degree from KAIST, and the Ph.D. in computer science from University of Maryland, College Park. He has been a Visiting Professor at KAIST and Sogang University in Korea, City University of Hong Kong, Ecole Centrale de Lille in France, and LinkopingUniversity and University of Skovde in Sweden

박 석

정보과학회논문지 : 데이타베이스
제 37 권 제 2 호 참조