

# 확률기반 실외환경 사고 인식기를 활용한 교통사고 추론 시스템

황주원<sup>○</sup> 이영설 조성배

연세대학교 컴퓨터과학과

[juwon@sclab.yonsei.ac.kr](mailto:juwon@sclab.yonsei.ac.kr), [tiras@sclab.yonsei.ac.kr](mailto:tiras@sclab.yonsei.ac.kr), [sbcho@cs.yonsei.ac.kr](mailto:sbcho@cs.yonsei.ac.kr)

## A Traffic Incident Inference System using Outdoor Environment Recognizer based on Probability

Ju-Won Hwang<sup>○</sup> Young-Seol Lee Sung-Bae Cho

Dept. of Computer Science, Yonsei University

### 1. 서론

최근 다양한 기능을 가진 디바이스와 센서의 발전으로 많은 분야에서 이를 이용하여 실세계의 문제를 해결하고 있다. 특히 상황인식 분야에서 변화가 많고 다양한 불확실성이 내재되어 있는 실외환경에서의 경우, 보다 풍부한 상황정보를 수집하기 위해서 많은 디바이스와 센서를 이용한다. 실제로 CCTV를 이용한 상황인식 시스템의 경우 특별한 기술 없이 모니터 요원들이 원격 감시하는 방식으로 가동되고 있다[1]. 이는 많은 인력소모와 모니터 요원의 감시실수와 같은 우를 범할 수 있으므로 효율적인 방법이 아니다. 따라서 이를 극복하기 위해 설치된 디바이스와 센서를 이용하여 수집한 데이터를 증거로 사건발생을 자율적으로 인지하는 시스템 개발이 요구되고 있다.

본 논문에서는 자율적으로 이벤트 발생을 감지하고 대응할 수 있는 시스템 개발을 목표로 하며, 인식하고자 하는 상황은 교차로에서의 차대차 사고와 차량단독사고를 포함한 차량 사고이다. 제안하는 방법은 정확하고 효율적으로 상황인식을 하기 위해 다음의 세단계로 구성하였다. 먼저, 사용가능한 로그 추출단계에서는 사고 인식에서 사용할 수 있는 가상 로그를 선별한다. 다음 단계에서는 추출한 로그들을 서버에서 사용 가능한 값으로 재정의 하였다. 또한 다양한 방법으로 모델링한 확률기반 실외환경 사고 인식기를 활용하여 도로의 혼잡한 상태가 사고인지, 신호에 의한 것인지 구분하고자 하였다. 마지막으로 사고 추론 단계에서는 로그를 입력하여 사고를 추론 및 인식하고자 본 논문에서 제안하는 사고 추론 시스템을 사용한다. 사고 추론 시스템은 앞서 설계한 인식기를 사용하여 사고와 로그의 관계 및 추론 과정을 시각화하며 다양한 방법을 이용하여 사고를 시뮬레이션 할 수 있는 것이 특징이다.

### 2. 본론

실제로 실외환경에서의 상태 인식이라 하면 고정되고 국한된 실내환경에서의 상태 인식보다 많은 상황을 고려해야 한다. 따라서 본 논문에서는 사고 인식구간에서 추출할 수 있는 가상 로그를 Speed, Background Sound, Traffic Density, Lanes Blocked, Flow, Alarm Whistle, Smoke, Signal 총 8개로 선별하여 전처리하였다. Speed, Background Sound, Traffic Density, Flow는 사고 인식 구간에서의 평균값을 사용하는 것으로, Background Sound는 도로의 소음 나타내는 decibel 단위 값으로 정의하였으며 Flow는 평균 교통량을 나타내는 것으로 신호에 따른 평균 교통량을 측정된 값을 나타내며 Traffic Signal은 추출한 로그가 교차로의 신호에 의한 값인지, 사고에 의한 값인지 구분하기 위한 것이다.

실외환경에서는 이벤트 발생과 관련하여 많은 증거들이 내재되어 있기 때문에 추출 가능한 로그와 사고 발생과의 관계를 정의한 후, 관계를 바탕으로 사고 발생 가능성을 판단해야한다. 이를 위해 인식기를 이용하여 앞서 전처리한 로그 값을 증거 데이터로 사용한다. 제안하는 방법에서는 불확실한 정보에서 보다 정확한 특징정보를 추출하기 위해 확률기반 연산을 하는 베이지안 네트워크를 이용하여 서로 다른 방법으로 사고를 추론하는 세 가지 종류의 사고 인식기를 설계하였다. 인식기 1은 사고를 인식하고자 하는 교차로에서 추출한 로그를 이용하여 사고 발생 여부를 추론한다. 이 경우 많은 로그를 사요하지 않아도 되므로 비교적 인식기 설계가 쉽다는 장점이 있지만 도메인을 단편적인 관점으로 단순화 시켰다는 문제가 있다. 이를 극복하고자 인식기 2와 인식기 3을 설계하였다. 인식기 2는 사고를 추론하고자 하는 교차로와 인접한 하나의 교차로에서 추출 가능한 로그를 이용한다. 두 교차로에서 추출한 로그의 흐름을 분석하는 방법으로, 로그를 비교하여 현재 교차로의 상태를 추론하는 것으로 사고를 인식하고자 하는 교차로에서 추출한 로그가 사고와 관련하여 특이성 있는 값이 아니더라도 사고인식이 가능하다는 것이 장점이다. 인식기 3은 사고를 인식하고자 하는 교차로에서 추출한 로그와 이전 시간대에서 발생한 사고 발생률을 증거데이터로 사용한다. 인식기 1과 2는 베이지안 네트워크를 이용하여 설계하였으며 인식기 3은 동적 베이지안 네트워크를 이용하여 설계하였다.

베이지안 네트워크는 중요한 변수들 간의 확률 관계를 노드와 호(arc)로 표현하는 방향성 비순환 그래프(DAG: directed acyclic graph) 형태이며[2], 노드마다 정의된 조건부 확률 테이블(CPT : conditional probability table)에

의해 적은 비용으로 노드간의 확률관계를 효율적으로 표현할 수 있는 모델이다. 이는 베이즈 규칙에 의해 조건부확률 연산이 가능하다. 수식 (1)은 조건부확률 연산 수식을 나타낸다. 이때,  $x_n$ 는  $n$ 번째 노드를,  $\pi_n$ 는  $x_n$ 의 부모 노드의 집합을 의미한다[3]. 동적 베이지안 네트워크는 일련의 확률변수들의 조건부확률분포를 표현하기 위한 확률 모델로서 특히 변수들 간의 시간적인 관계를 표현하는데 유용한 모델이다. 일반적으로 동적 베이지안 네트워크 모델에서는 마르코프특성에 의해 현재 시점의 확률변수는 이전 N단계까지의 영향을 받고 그 이전의 시간단계에서는 영향을 받지 않는다. 동적베이지안 네트워크의 식은 수식 (2)와 같다[4].

$$\begin{aligned}
 P(x_1, x_2, \dots, x_n) &= P(x_1, x_2, \dots, x_{n-1})P(x_n | x_{n-1}, x_{n-2}, \dots, x_1) \\
 &= P(x_1)P(x_2 | x_1)P(x_3 | x_2, x_1)\dots \\
 &= P(x_1)P(x_2 | \pi_2)P(x_3 | \pi_3)\dots P(x_n | \pi_n)
 \end{aligned}
 \quad (1)$$

$$P(Z_1, \dots, Z_T) = \prod_{t=1}^T \prod_{i=1}^N P(Z_t^i | Pa(Z_t^i)) \quad (2)$$

베이지안 네트워크를 이용하여 설계한 모델의 복잡성은 CPT 파라미터 수를 이용하여 평가한다. 결과노드로 많은 원인노드들이 연결되면 CPT 파라미터 수가 기하급수적으로 증가해 복잡성이 커진다. 때문에 앞서 설계한 인식기의 경우 사고노드를 원인노드로, 증거노드를 결과노드로 설계하였다. 이러한 설계방법은 베이지안 네트워크는 역추론을 할 수 있기 때문에 가능하다. 인식기 1의 경우 CPT 파라미터수가 총 72개로, 사고 노드를 결과노드로 설계하였을 때의 CPT 파라미터 수인 총 208보다 약 3분의 1로 줄일 수 있다.

본 논문에서 제안하는 사고 추론 시스템은 앞서 설계한 세가지 인식기를 이용하여 사고를 시뮬레이션 할 수 있으며 추출한 로그를 로드하여 사고를 인식할 수 있다는 것이 특징이다. 시스템을 이용하여 가상 데이터에 대한 임의의 값을 입력하여 사고를 추론할 수 있게 하였으며, 실제 데이터를 로드하여 사고인식을 할 수 있게 하였다. 또한, 인식기를 활용한 추론 및 인식과정, 추론된 확률값과 결과를 시각화하였다. 본 시스템은 사용자가 데이터 입력에 의한 직접 추론 및 추론 과정을 시뮬레이션할 수 있도록 하였다.

사고 인식기의 성능을 테스트하기 위하여 가상시나리오를 사용하였다. 가상시나리오는 신호에 의한 혼잡한 상태의 시나리오와 사고에 의한 혼잡한 상태의 시나리오로 구성하였다. 인식기 1은 12개의 시나리오를 사용하였고 Detection Rate는 100%, False alarm Rate는 28%, 인식기 2는 14개의 시나리오를 사용하여 Detection Rate 82%, False alarm Rate 33%, 인식기 3은 24개의 시나리오를 사용하여 Detection Rate 100%, False alarm Rate 33%의 성능을 보임을 확인하였다. 인식기 3의 경우 상황인식, 교통흐름저하, 사고정리 상황의 추론 값이 다른 인식기와 비교하여 높음을 확인할 수 있었다. 이는 이전 시간대에 사고가 발생하였음을 증거 값으로 사용하였기 때문이다.

### 3. 결론

본 논문에서는 교차로에서의 사고 발생을 자율적으로 감지하고 대응하기 위해 확률기반 사고 인식기를 활용하였고 사고 추론 과정과 추론 및 인식 과정을 시각화하기 위해 사고 추론 시스템을 제안하였다. 제안하는 사고 추론 시스템은 사고 인식을 위해 사용가능한 로그 추출, 서버에서의 로그 분석 및 사고 인식, 가상 데이터 및 실제 데이터를 이용한 사고 추론의 네 단계로 구성하였다. 사고 인식단계에서는 베이지안 네트워크로 모델링한 인식기를 사용하였다. 베이지안 네트워크는 불확실한 실외환경에서 추출한 로그를 이용한 추론에 유용하기 때문이다. 데이터를 이용한 사고 추론 단계에서는 제안하는 시스템을 이용하여 추론 과정 및 추론 결과를 시각화하였고 시뮬레이션이 가능하게 하였다.

향후 연구는 가상 시나리오가 아닌 실제 데이터 및 교통사고 시뮬레이터 기반 데이터를 사용하여야 할 것이며 차대차가 아닌 차대사람 등 발생 가능한 사고로 확장하여 추론 및 인식할 수 있도록 해야 할 것이다.

### 감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업(NIPA-2009-(C1090-0902-0046))과 지식경제부 및 한국산업기술평가관리원의 산업원천기술개발사업(10033776, 대규모 지능형 협업 무인감시 시스템 원천 기술개발)의 연구결과로 수행되었음.

### 참고문헌

- [1] <http://www.aynews.net/>
- [2] D. Heckerman, Bayesian networks for data mining," Data Mining and Knowledge Discovery, vol. 1, no. 1, pp. 79-119, 1997.
- [3] K. B. Korb and A. E. Nicholson, Bayesian Artificial Intelligence, Chapman & hall/CRC, 2003.
- [4] H. S. Park and S. B. Cho, "Prediction of user activity based on mobile life-log using dynamic Bayesian network," Proc. the 30th KIPS Fall Conference, vol. 15, no. 2, pp. 60-63, 2008.