■ 연구논문 요약문1

논문제목	Representation learning for unsupervised heterogeneous multivariate time series segmentation and its application
게재정보	Computers & Industrial Engineering, vol130, 2019
개요	- 연구 목적과 내용에 대해 소개 (서술식으로 작성해도 됨) - While driving a vehicle, data are collected from a huge number of sensors that generate both categorical and continuous variables with varying scales. In order to understand the status of the vehicles and the drivers' behaviors, it is crucial to segment and identify different phases within this time series data. However, data often lacks labels to denote different phases, rendering supervised learning based segmentation methods as futile. - Consequently, distance based time series segmentation method is a realistic solution for detecting different phases in the sensor data. However, there is no universal distance measure that utilizes both categorical and continuous variables simultaneously to segment the multivariate data.
연구결과	- 연구개발 결과를 가능한 한 비전문가가 이해할 수 있도록 설명 (서술식으로 작성해도 됨) - - In this paper, we propose a novel unsupervised time series segmentation framework for heterogeneous multivariate data. By applying the distributed representation of the word embedding methods, we transform multivariate heterogeneous data into continuous vectors, allowing them to be segmented by conventional distance metrics such as Euclidean or Cosine distance. Subsequently, similar segments are clustered to generate general patterns. Without any labels or feature engineering, our framework successfully segments and discovers insightful driving patterns from heterogeneous sensor data collected from actual vehicles.
활용분야 및 기대효과	- 연구개발 결과의 활용 분야와 기대 효과를 서술 (서술식으로 작성해도 됨) - Real time stream data, including automobile sensor data, is better prepared with the approach proposed in the paper. - Which will lead to better processing and better understanding of what is going on in the system of interest. - Better maintenance and better early warning is all possible.