Accepted Manuscript

Big Data for Context Aware Computing – Perspectives and Challenges

Kalyan P. Subbu, Athanasios V. Vasilakos

Please cite this article in press as: K.P. Subbu, A.V. Vasilakos, Big Data for Context Aware Computing – Perspectives and Challenges, *Big Data Res.* (2017), https://doi.org/10.1016/j.bdr.2017.10.002

Big Data for Context Aware Computing - Perspectives and Challenges

Kalyan P. Subbu,   Athanasios V. Vasilakos, Senior Member IEEE,
[pathapks@gmail.com,th.vasilakos@gmail.com]

Abstract

Big data has arrived. Myriad applications, systems generate data of humongous volumes, variety and velocity which traditional computing systems and databases are unable to manage. The proliferation of sensors in every possible device is also becoming one of the major generators of Big data. Of particular interest in this article is how context aware computing systems which derive context from data and act accordingly, deal with such huge amounts of data. Big industry players namely Google, Yahoo, and Amazon are already developing context aware applications using user data from emails, chat messages, browsing and shopping histories etc., For instance, Gmail reminds us of our flight schedule by understanding flight booking related content in our emails. Similarly, Amazon understands user preference and recommends items of interest to shop and so on. In this paper, we survey context aware computing systems from a Big data perspective. We first propose a taxonomy of existing work on the basis of sensing platforms and then discuss the latest developments in this field of Big data context aware systems focussing on how such systems deal with various Big data challenges. We conclude the paper with an insight on open research issues involving designing and developing context aware Big data generating systems.

Keywords: Big data, Context Awareness, Smartphones, Sensors, Internet of Things, Wireless Sensor Networks.

1. Introduction

The recent years has witnessed data flow from a multitude of sources. This led to the Big Data revolution. Big data characteristics include a) volume (amount generated), b) velocity (rate of generation), c) variety (originate from various sources) and d) volatality(uncertain and noisy data). A major source of Big data during the last decade originated from Internet related applications such as user emails, multimedia (audio and video), online transactions, shopping and browsing preferences and the Internet of Things (IoT). Other venues identified include scientific data (astronomy), medical (health data records, patient data), biomedical (DNA sequencing), transportation(traffic, routes, taxi demands), entertainment (Netflix), and recommendation systems(travel, music, shopping).

Another domain of interest in this article is Context aware computing (CAC) which enables applications to have an awareness of the context by making inferences from the data collected and providing smart intelligent services to the user. Data sensed is typically used to extract some information about a context, which can refer to 1) Computing/communication context (network connectivity, communication costs, resource accessibility), 2) User context (user profile, location, activity, social situation, preference), 3) Physical context (lighting, temperature, noise, traffic conditions), or 4) Time context (hour of day, day of week, season, year). Domains such as mobile and pervasive computing are two areas which have seen umpteen number of innovative systems with context awareness. Wireless sensor networks (WSNs) a field under mobile computing paved the way for sensing and computation on battery powered devices of small size. However, WSNs lacked the pervasive/ubiquitous property, meaning facilitating sensing and computation anywhere and everywhere. The development of micro eletro-mechanical sensors (MEMS) sensors in devices particularly the mobile phone started an era of pervasive/mobile sensing which was coupled with unobtrusive, continual, and reliable connectivity. Therefore, in contrast to distributed and mobile computing, pervasive computing extends sensing, computing and communication capabilities to the next level by integrating with