

Scalable Uncertainty-Aware Truth Discovery in Big Data Social Sensing Applications for Cyber-Physical Systems

ABSTRACT

Social sensing is a new big data application paradigm for Cyber-Physical Systems (CPS), where a group of individuals volunteer (or are recruited) to report measurements or observations about the physical world at scale. Discovering the correctness of reported observations and reliability of data sources without prior knowledge on either of them. We refer to this problem as truth discovery. While prior studies have made progress on addressing this challenge, two important limitations exist: current solutions did not fully explore the uncertainty aspect of human reported data, which leads to sub-optimal truth discovery results; current truth discovery solutions are mostly designed as sequential algorithms that do not scale well to large-scale social sensing events. In this paper, we develop a Scalable Uncertainty-Aware Truth Discovery (SUTD) scheme to address the above limitations. The SUTD scheme solves a constraint estimation problem to jointly estimate the correctness of reported data and the reliability of data sources while explicitly considering the uncertainty on the reported data. To address the scalability challenge, the SUTD is designed to run on a Graphic Processing Unit (GPU) with thousands of cores, which is shown to run two to three orders of magnitude faster than the sequential truth discovery solutions.

EXISTING SYSTEM

In existing, a fundamental challenge in social sensing applications lies in discovering the correctness of reported observations and reliability of data sources without prior knowledge on either of them. We refer to this problem as truth discovery. While prior studies have made progress on addressing this challenge, two important limitations exist: current solutions did not fully explore the uncertainty aspect of human reported data, which leads to sub-optimal truth discovery results; current truth discovery solutions are mostly designed as sequential algorithms that do not scale well to large-scale social sensing events.

DRAWBACKS

- Current solutions did not fully explore the uncertainty aspect of human reported data, which leads to sub-optimal truth discovery
- Current truth discovery solutions are mostly designed as sequential algorithms that do not scale well to large-scale social sensing events.

PROPOSED SYSTEM

In this paper, we develop a Scalable Uncertainty-Aware Truth Discovery (SUTD) scheme to address the above limitations. The SUTD scheme solves a constraint estimation problem to jointly estimate the correctness of reported data and the reliability of data sources while explicitly considering the uncertainty on the reported data. To address the scalability challenge, the SUTD is designed to run a Graphic Processing Unit (GPU) with thousands of cores, which is shown to run two to three orders of magnitude faster than the sequential truth discovery solutions. In evaluation, we compare our SUTD scheme to the state of the art solutions using three real world datasets collected from Twitter:.

ADVANTAGES

- It solves a constraint estimation problem to estimate both the correctness of reported data and the reliability of data sources.
- It improves both the estimation accuracy and execution time of current truth discovery solutions.

SYSTEM REQUIREMENTS

H/W System Configuration:-

- Processor - Pentium –IV
- RAM - 4 GB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

S/W System Configuration:-

- Operating System : Linux
- Application Server : Tomcat5.0/6.X
- Backend coding : Java
- Tool : Virtual Box
- Environment : Ubuntu
- Technology : Hadoop