Automated Anomaly Detection
Master's Project in Software Engineering

About WideOrbit:

Since 1999, WideOrbit is the world's leading provider of advertising management technology for cable networks, local television stations and radio stations. More than 3,200 broadcasters and networks leverage WideOrbit solutions to streamline operations, maximize revenue from traditional, digital and programmatic channels, and extend their business across distribution platforms. WideOrbit is headquartered in San Francisco with offices across the United States as well as in London, Paris and Gothenburg, Sweden.

WideOrbit Sweden (formerly Admeta) is fully owned subsidiary of WideOrbit. We are a leading provider of solutions for digital programmatic sell-side platforms for internet advertising. This means that we provide software to auction internet ad impressions. We optimize each impression for maximum revenue using sophisticated optimization technologies. We are currently processing hundreds of thousands web request per second generating terabytes of data daily, showing tens of billions of ad impressions monthly. This traffic is increasing rapidly which puts new requirements on our constantly evolving system.

Thesis Outline:

The high throughput and the real-time nature of our system call for real-time monitoring. Currently we use two third-party providers for monitoring health of our systems, as well as in-house alarm system covering only the most crucial operations. The output of these is then aggregated through another third-party vendor and checked manually. The current alarm system checks if a handful of numeric indicators manually picked from the whole data array exceed manually set thresholds. There is a need for a smarter/more automated monitoring system capable of proposing data-driven criteria for catching outliers and using them to detect anomalies in the system.

The goal of this thesis is to develop a framework for automatic detection of anomalies across multiple metrics of various kinds aggregated from the many sources of diagnostic data. The framework should be able to:

- fetch and aggregate the metrics from multiple sources (MySQL database, 3rd party API's),
- summarize and visualize the many sources of data,
- detect anomalies/outliers across several different metrics,
- automatically emphasize the most relevant metrics at the moment

Backend work will preferably be done using python and/or C#/.NET, but other tools and languages can be considered.

To Apply:
Contact David Burke: dburke@wideorbit.com