DOCTORAL CANDIDATE:	Vinay Jayarama Setty
DEGREE:	Philosophiae Doctor
FACULTY:	The Faculty of Mathematics and Natural Sciences
DEPARTMENT:	Department of Informatics
AREA OF EXPERTISE:	Distributed Systems, Publish/Subscribe Systems, Social Networks, Algorithms
SUPERVISORS:	Roman Vitenberg, Maarten van Steen
DATE OF DISPUTATION:	18th of February 2015

DISSERTATION TITLE: ***Publish/Subscribe for Large-Scale Social Interaction: Design, Analysis and Resource Provisioning***

Scientific Summary (English):

With the widespread use of online social networking, we are witnessing an information explosion due to tremendous amount of messages exchanged between users. These messages are often delivered as notifications to the user devices such as smartphones, tablets and laptops. Publish/subscribe (pub/sub) is a communication paradigm, which is widely used for delivering such notifications. The pub/sub paradigm is typically used for RSS feed notifications, financial data dissemination and business process management. Additionally, the music streaming service of Spotify utilizes pub/sub paradigm to deliver the notifications generated due to the user activity, artist updates etc. Such notifications have grown exponentially in recent years to the order of billions of notifications generated by the millions of active users every day. However, the existing pub/sub systems are not yet ready to support such a scale: They need to address several challenges first.

In this dissertation, first we extensively studied an existing pub/sub system—the Spotify pub/sub. That enabled us to identify several research challenges that need to be addressed for scaling the pub/sub systems. One of the interesting observations we made was that social notifications often overwhelm the users and they tend to ignore them when the number of delivered notifications grows beyond a certain threshold. Exploiting this observation, we proposed a new pub/sub architecture that avoids delivering notifications beyond a certain threshold to the users. By doing this we were able to reduce the load on the servers driving the pub/sub services. We further explored the possibility deploying such a pub/sub service in cloud environments. We identified the costs involved in deploying pub/sub in the cloud and we provided cost-saving techniques by organizing the data in a smart way.

In the process of designing a pub/sub service described above, we identified several novel problems. We also proposed several novel algorithms and heuristics to solve those problems. In order to validate their efficiency and correctness, we simulated the pub/sub environment using data collected from the real pub/sub systems such as Spotify. In the results we showed that our techniques lead to efficient, scalable, and cost-effective deployment. In addition to the empirical study, we also performed formal analysis in order to gain an understanding of the hardness of the problems and quality guarantees offered by the heuristics and algorithms we proposed. The findings of the research conducted in this dissertation have been published in several reputed international computer science conferences such as ACM/IFIP/USENIX Middleware conference, ACM Distributed Event Based Systems (DEBS), IEEE International Conference on Computer Communications (Infocom) and IEEE International Conference on Distributed Computing Systems (ICDCS).

Scientific Summary (Norwegian):

Det vil i fremtiden ble svært viktig å kunne utvikle skalerbare arkitekturer som kan håndtere den økende datatrafikken i sosiale nettverk. Et eksempel vi ser på i denne avhandlingen er musikkstrømmingstjenesten Spotify, som i tillegg støtter sosial interaksjon mellom brukere og musikkartister. Gjennom denne tjenesten kan Spotify levere korte beskjeder til brukere om nye musikkalbum, felles spillelister og aktiviteter som venner driver med. Allerede i dag leverer Spotify flere milliarder slike beskjeder hver eneste dag. Andre eksempler omfatter mange milliarder sosiale beskjeder i Facebook og Twitter. Databehandling i en slik skala krever betydelige resurser samt skalerbare implementasjoner.

I denne avhandlingen utvikler vi nye metoder for å kjøre slike tjenester i skyer (clouds) og datasentre, og avhandlingen diskuterer resurstilordninger og kostnadredusering. Vi validerer metodene som foreslåes gjennom simuleringer og formell analyse. Resultatene i denne avhandlingen er publiserte i viktige internasjonale konferanser.