

Operating Complex IT-Systems Seminar – Collection of Topics

Winter term 2016/17

This document lists the seminar topics selected by our research staff for this term’s bachelor and master seminars. Each student is encouraged to choose at least 3 desired topics from the provided set. Each topic includes an abstract, references and a recommendation regarding the applicability for bachelor or master students (or possibly both). The references are to be used as a foundation for the literature review, which is required to prepare the slides and the scientific paper. Note that it is not sufficient to rely on the provided references only. If you have trouble accessing the references (e.g. through ieeexplore) use a workstation that is either physically or virtually (e.g. by VPN) connected to the university network. You may additionally try to locate a free-access version of the reference using the scholar search engine offered by Google.

Further notes:

- Please make sure not to miss the servicetalk on scientific writing and presentation on 10.11.2016 16:00-18:00 (see website for details)
- Final assignment of the topics takes place during the kickoff meetings on 19.10.2016 10:00-12:00 (see website for details). Participation is mandatory. Students not showing up unexcused automatically cancel their registration.
- Please select at least three topics of interest. We will try to resolve conflicts, but we can’t guarantee that each student will be assigned his favorite topic.

Topics sorted by area of research

1	Distributed Systems Theory	1
2	Large-Scale Cluster Computing.....	4
3	Big Data Analytics & Visualization	7
4	Internet of Things & Machine-2-Machine Communication	11
5	Miscellaneous Topics.....	14

1 Distributed Systems Theory

Topic 1: Virtual Time
Abstract: Virtual time is a new paradigm for organizing and synchronizing distributed systems which can be applied to such problems as distributed discrete event simulation and distributed database concurrency control. Virtual time provides a flexible abstraction of real time in much the same way that virtual memory provides an abstraction of real memory. It is implemented using the Time Warp mechanism, a synchronization protocol distinguished by its reliance on lookahead-rollback, and by its implementation of rollback via antimessages.

References:	
http://dl.acm.org/citation.cfm?id=3988	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 2: Optimistic Concurrency Control	
Abstract:	
<p>Most approaches to concurrency control in database systems rely on locking of data objects as a control mechanism. In this paper, two families of nonlocking concurrency controls are presented. The methods used are “optimistic” in the sense that they rely mainly on transaction backup as a control mechanism, “hoping” that conflicts between transactions will not occur. Applications for which these methods should be more efficient than locking are discussed.</p>	
References:	
[1] http://dl.acm.org/citation.cfm?id=319567	
[2] http://redis.io/topics/transactions	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 3: Software Transactional Memory	
Abstract:	
<p>In computer science, software transactional memory (STM) is a concurrency control mechanism analogous to database transactions for controlling access to shared memory in concurrent computing. It is an alternative to lock-based synchronization. STM is strategy implemented in software, rather than as a hardware component. A transaction in this context occurs when a piece of code executes a series of reads and writes to shared memory. These reads and writes logically occur at a single instant in time; intermediate states are not visible to other (successful) transactions.</p>	
References:	
[1] http://groups.csail.mit.edu/tds/papers/Shavit/ShavitTouitou.pdf	
[2] http://dl.acm.org/citation.cfm?id=872048	
[3] http://blog.enfranchisedmind.com/2009/01/the-problem-with-stm-your-languages-still-suck/	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 4: Differential Synchronization	
Abstract:	
<p>Differential Synchronization (DS) method is for keeping documents synchronized. The key feature of DS is that it is simple and well suited for use in both novel and existing state-based applications without requiring application redesign. DS uses deltas to make efficient use of bandwidth, and is fault-tolerant, allowing copies to converge in spite of occasional errors. We consider practical implementation of DS and describe some techniques to improve its performance in a browser environment.</p>	
References:	
[1] https://neil.fraser.name/writing/sync/eng047-fraser.pdf	
[2] https://spring.io/blog/2014/10/22/introducing-spring-sync	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 5: Geo Replication

Abstract:	
Geo-replication systems are designed to improve the distribution of data across geographically distributed data networks. This is intended to improve the response time for applications such as web portals. Geo-replication can be achieved using software, hardware or a combination of the two. Online services distribute and replicate state across geographically diverse data centers and direct user requests to the closest or least loaded site. While effectively ensuring low latency responses, this approach is at odds with maintaining cross-site consistency.	
References:	
[1] https://www.usenix.org/system/files/conference/osdi12/osdi12-final-162.pdf	
[2] https://www.usenix.org/system/files/conference/osdi14/osdi14-paper-ardekani.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 6: CoRAL – Reliable Web Services	
Abstract:	
Making stateful web services reliable requires elaborate cross-layer techniques. The fault tolerance scheme CoRAL (Connection Replication and Application-level Logging) actively replicates the state of a TCP connection and additionally logs HTTP requests/replies to enable fast failover to a warm-standby server.	
References:	
http://web.cs.ucla.edu/csd/research/labs/csl/projects/coral/	
http://web.cs.ucla.edu/~tamir/papers/pdcs03.pdf	
http://millennium.cs.ucla.edu/~tamir/papers/coral_jss09.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 7: Distributed State Machines	
Abstract:	
Maintaining consistent application state is an important issue when implementing replicated network services. Paxos is a widely used algorithm for implementing a Distributed State Machine which allows a number of service replicas to maintain consistency. Paxos has been extended and improved many times since Lamports original description of the algorithm.	
References:	
Paxos	
http://dl.acm.org/citation.cfm?doid=279227.279229	
http://www.ux.uis.no/~meling/papers/2013-paxostutorial-opodis.pdf	
NetPaxos	
NetPaxos is an extension to Paxos optimizing it for the use in modern SDN-capable switches.	
http://perso.uclouvain.be/marco.canini/papers/netpaxos.sosr15.pdf	
Raft	
Raft is a novel consensus algorithm inspired by Paxos designed to be more understandable for students of dependable systems while providing the same consistency guarantees and performance as Paxos.	
http://www.eecs.harvard.edu/cs261/papers/ongaro14.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: Besides the original Paxos algorithm, this topic contains 2 sub-topics dealing with extended approaches. Students can decide whether to present one topic in detail or focus on a comparison. Furthermore, this topic may be assigned to up to three students as well.	

Topic 8: Partial Synchronous Distributed Systems

Abstract:	
System models are an important tool to understand the properties and the behavior a Distributed System and its communicating nodes exhibit. Commonly known models are the synchronous and the asynchronous model. Efficient algorithms for most of the problems we face in Distributed Systems are known for the synchronous model. However, it is almost impossible to implement the synchronous model in real life. The asynchronous model instead, rather characterizes real world systems like the Internet. Unfortunately, the assumptions of this model are too weak to solve several important problems efficiently. As a remedy, the partial synchronous model is introduced.	
References:	
http://www-usr.inf.ufsm.br/~ceretta/papers/MITLCSTM270.pdf https://ecommons.cornell.edu/bitstream/handle/1813/7192/95-1535.pdf?sequence=1	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

2 Large-Scale Cluster Computing

Topic 9: Publish/Subscribe with Apache Kafka	
Abstract:	
In modern distributed systems the communication between components often relies on messaging. Kafka is a distributed messaging system that was developed for collecting and delivering high volumes of log data with low latency. A few unconventional yet practical design choices in Kafka make it efficient and scalable with performance superior to popular alternatives.	
References:	
http://notes.stephenholiday.com/Kafka.pdf , https://www.rabbitmq.com/	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 10: Fault-Tolerance in MapReduce	
Abstract:	
MapReduce is an often used Parallel programming Framework. It enables the user to run jobs within the cloud on many machines. The usage of a lot of working nodes, leads to an increase of failure probability. Therefore it is necessary to work on fault tolerance for MapReduce Systems. There are several ideas how to improve fault tolerance in MapReduce:	
References:	
Improving MapReduce fault tolerance in the cloud http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5470865&	
Towards self-caring mapreduce: Proactively reducing fault-induced execution-time penalties http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5999808	
HPCLOUD: A novel fault tolerant architectural model for hierarchical MapReduce http://ieeexplore.ieee.org/mobile/Abstract.jsp?arnumber=6206822	
http://ieeexplore.ieee.org/document/7152978/	
BeTL: MapReduce Checkpoint Tactics Beneath the Task Level http://ieeexplore.ieee.org/document/7004380/	
ReCT: Improving MapReduce performance under failures with resilient checkpointing tactics	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 11: Comparison of Distributed File Systems for Parallel Execution	
Abstract:	
Running jobs in the cloud in parallel raises the needs of storing data efficiently. Usually parallel execution Frameworks like MapReduce use distributed File systems. For example HDFS, Google FS, CODA, GlusterFS...	
But which one should be used, in which case?	
References:	
HDFS http://pristinespringsangus.com/hadoop/docs/hdfs_design.pdf	
GFS http://dl.acm.org/citation.cfm?id=945450	
CODA http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=54838	
Applicable for BSc: yes	Applicable for MSc: no
Further Notes: This topic is related to an available/vacant bachelor thesis. In case you wish to use your seminar results as a foundation for a BA thesis literature review, contact Mareike Höger for further details.	

Topic 12: Dominant Resource Fairness: Fair Allocation of Multiple Resource Types	
Abstract:	
Fairness is one of the key requirements when users share resources such as CPUs, memory, and I/O. Dominant Resource Fairness is an approach to provide fairness across multiple resources at the same time. It is used with the two popular open-source resource management systems for large-scale data analysis: Apache Hadoop (YARN) and Apache Mesos.	
References:	
http://static.usenix.org/events/nsdi11/tech/full_papers/Ghodsi.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 13: Automatic Resource Provisioning for Data-parallel Processing Systems	
Abstract:	
Data-parallel processing frameworks like MapReduce, Flink, and Spark arguably make analysis of very large datasets easier. Users create programs using a small set of pre-defined operators and write sequential code to configure these. The frameworks manage task parallelization and distribution as well as handle node failures. However, users do still need to specify how much and which resources to use for their jobs. This is often difficult and users consequently tend to overprovision significantly to ensure minimal performance requirements.	
Addressing this problem there are multiple systems such as Ernest and Jockey that predict the runtime of jobs and then select resources according to users' runtime targets. This allows users to specify their actual performance goals instead of having to guess an adequate set of resources.	
References:	
https://amplab.cs.berkeley.edu/publication/ernest-efficient-performance-prediction-for-large-scale-advanced-analytics/	
http://research.microsoft.com/en-us/UM/people/srikanth/data/jockey_eurosys12.pdf	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 14: Iterative Parallel Dataflows

Abstract:	
<p>Many important algorithms are iterative. These include, for example, many algorithms for graph analysis and machine learning. Parallel dataflows enable users in analyzing large datasets using clusters of computers. Data is processed through a graph of operators such as Map, Reduce, Join and GroupBy. These are executed in parallel and across many nodes.</p> <p>With iterative algorithms the same parallel dataflow is executed repeatedly. This fact can be utilized in various ways, which is the theme of this topic.</p>	
References:	
https://cs.stanford.edu/~matei/papers/2012/nsdi_spark.pdf http://stratosphere.eu/assets/papers/spinningFastIterativeDataFlows_12.pdf http://www.vldb.org/pvldb/vol6/p1678-popescu.pdf	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 15: Making Sense of Performance in Data Analytics Frameworks	
Abstract:	
<p>Making informed decisions in the design and implementation of large-scale data processing frameworks requires a good understanding of their performance, including knowing which resources are the bottleneck and therefore determine the performance of applications. Blocked Time Analysis is an approach to identify and quantify such bottlenecks for distributed processing frameworks. Knowing this helps to understand the performance of specific workloads using a specific frameworks and clusters, yet deriving facts about entire classes of systems—including separating conceptual reasons from implementation reasons—is still difficult.</p>	
References:	
https://amplab.cs.berkeley.edu/publication/making-sense-of-performance-in-data-analytics-frameworks/ http://www.frankmcherry.org/pagerank/distributed/performance/2015/07/08/pagerank.html http://www.frankmcherry.org/assets/COST.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 16: SQL on Hadoop	
Abstract:	
<p>Big data applications require a mix of processing techniques, data sources and storage formats. The earliest systems designed for these workloads, such as MapReduce, gave users a powerful, but low-level, procedural programming interface. Programming such systems was onerous and required manual optimization by the user to achieve high performance. As a result, multiple new systems (SparkSQL, Hive, Impala) sought to provide a more productive user experience by offering relational interfaces (e.g. SQL) to big data. The objective of this topic is to compare these systems with each other by.</p>	
References:	
https://amplab.cs.berkeley.edu/publication/spark-sql-relational-data-processing-in-spark/ https://hive.apache.org/ http://cidrdb.org/cidr2015/Papers/CIDR15_Paper28.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 17: Tiered Storage on Hadoop	
Abstract:	

One Big Data main challenge is to deal with the deal with the exponentially growing data volumes, and to do so in an economically viable fashion. A promising trend in storage technologies is the emergence of heterogeneous and hybrid storage systems that deploy different types of storage devices, e.g. SSDs, HDDs, and ramdisks. The objective of this topic is to analyze recent developments for HDFS to support hybrid and tiered storage systems.

References:

<http://www.ebaytechblog.com/2015/01/12/hdfs-storage-efficiency-using-tiered-storage/>
<http://people.cs.vt.edu/butta/docs/ccgrid2014-hats.pdf>
<http://pages.cs.wisc.edu/~akella/CS838/F15/838-CloudPapers/hdfs.pdf>
<http://dl.acm.org/citation.cfm?id=2670985>
<http://www.alluxio.org/>

Applicable for BSc: yes

Applicable for MSc: yes

Further Notes: no

3 Big Data Analytics & Visualization

Topic 18: IBM Watson – AlchemyData News

Abstract:

IBM developed a questioning answering computer called Watson, which won the quiz show Jeopardy in 2011. Since then, the IBM Watson team is growing and produces a lot of different intelligent services. One of them is called AlchemyData News, which uses Natural Language Processing and further Text Mining techniques in order so annotate given news articles. Such annotations are further used to build intelligent services upon the news data by third party organizations.

References:

<http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/alchemy-data-news.html> ,
<http://querybuilder.alchemyapi.com/builder> ,
<http://resources.alchemyapi.com/white-papers>

Applicable for BSc: yes

Applicable for MSc: no

Further Notes: no

Topic 19: Ontology learning from natural texts

Abstract:

Retrieving knowledge from natural texts is a hard task for machines, while humans manage to detect relations between words easily. In order to empower machines to interact with knowledge, the concept of ontologies was introduced. Such automatic interaction with knowledge can be further used to enable building advanced recommendation engines and give machines the ability to reason about worlds information. As todays big data content is created by users in a textual way, techniques are needed to transfer most important information from a given natural text to an ontology.

References:

<http://ontotext.com/> ,
https://www.researchgate.net/publication/29650100_Text-based_ontology_construction_using_relational_concept_analysis ,
<http://wortschatz.uni-leipzig.de/~cbiemann/pub/2005/BiemannLDVOntology05.pdf> ,
<http://userpages.uni-koblenz.de/~staab/Research/Publications/2009/handbookEdition2/ontology-learning-handbook2.pdf> ,
<http://www.semanticsoftware.info/system/files/lrec2010-owlexporterV2.pdf> ,
<http://olp.dfki.de/OntoLT/esws2004.pdf>

Applicable for BSc: yes

Applicable for MSc: yes

Further Notes: no

Topic 20: Folksonomies – Crowded based Ontologies	
Abstract: In social media, tagging is a well known concept to mark cross linked microblogs talking about similar topics. These tags can be used to populate knowledge bases called Folksonomies, also known as crowded based ontologies. In addition, Folksonomies can be further used to create powerful recommender systems.	
References: https://www.researchgate.net/profile/Dominik_Kowald/publication/280920102_Evaluating_Tag_Recommender_Algorithms_in_Real-World_Folksonomies_A_Comparative_Study/links/55cb245308aebc967dfc05d9.pdf , http://www.gartner.com/it-glossary/folksonomy	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 21: Distributed Topic Modeling Algorithm	
Abstract: Topic modeling is a type of statistical model to detect “topics” in a set of given documents. Performing learning of topic models from a massive corpus of possibly millions of documents raises the need of scalable learning approaches. This topic shows a scalable asynchronous distributed algorithm for topic model learning.	
References: http://arxiv.org/pdf/1412.4986.pdf	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 22: Visualizing Topic Models	
Abstract: Topic modeling is a type of statistical model to detect “topics” in a set of given documents. This approach tackles the problem of document summarization besides various problems. In order to give humans the ability to interact with such topic models, several techniques are proposed in order to handle the information overload.	
References: http://inphodata.cogs.indiana.edu/oldbailey/40/papers/aaai15-topic-explorer-demo.pdf , http://www.ics.uci.edu/~asuncion/pubs/TIST_11.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 23: Cleaning up messy data using Protégé and Open Refine	
Abstract: Today’s traditional relational databases are build for a single use of purpose, such migration of data from different databases can cause conflicts in choosing the right new database structure. In order to overcome this kind of problem, Google introduced a project named Refine. Refine uses techniques to easily structure data and modifies not wanted entries. Furthermore, Protégé can be used to build an ontology, which represents the RDBS stored data. The ontology can be easily integrated to the given dataset using Open Refine.	
References:	

http://protege.stanford.edu/ , http://openrefine.org/	
Applicable for BSc: yes	Applicable for MSc: no
Further Notes: no	

Topic 24: Decision Tree Online Learning	
Abstract:	
<p>A lot of data is retrieved in a streaming online scenario setting. Such information streams need to be evaluated in real-time by machines. The real-time constrained can be succeeded by using decision trees, which need to be learned in a previous step from a given data stream. Furthermore, in an online scenario, concept shift is a known problem, which needs to be coped with when building a learning model from such data streams to enable a real-time evaluation.</p>	
References:	
<p>http://www.cs.princeton.edu/courses/archive/spr07/cos424/papers/mitchell-dectrees.pdf, https://www.researchgate.net/profile/Geoffrey_Holmes3/publication/225395781_Fast_Perceptro_n_Decision_Tree_Learning_from_Evolving_Data_Streams/links/00b7d5159497fdff17000000.pdf , ftp://ftp7.freebsd.org/sites/ftp.sourceforge.net/pub/sourceforge/m/mo/moa-datastream/documentation/Manual.pdf</p>	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 25: Facebook AI	
Abstract:	
<p>Facebook exploits new market opportunities created from its massive social media content. Therefore, Facebook introduced a department “Facebook AI Research” (FAIR). Besides Facebooks outstanding face recognition service and social graph API, FAIR introduced recently its Messenger Bots. Such services enable automatic personalized social communication. In order to understand the vision of Facebook, this topic describes Facebook’s recent technologies and research topics.</p>	
References:	
<p>https://developers.facebook.com/docs/messenger-platform, http://newsroom.fb.com/news/2016/04/using-artificial-intelligence-to-help-blind-people-see-facebook/, https://research.facebook.com/publications/how-blind-people-interact-with-visual-content-on-social-networking-services/, https://code.facebook.com/posts/457605107772545, https://research.facebook.com/blog/fair-open-sources-deep-learning-modules-for-torch/, https://wit.ai/</p>	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 26: Image Recognition – Cloud based Services	
Abstract:	
<p>The social media content of digital images and videos is growing in recent years. Such content contains a lot of information, hidden in it’s data. As machines read only binary data, objects on an image are hidden to the machine. In order to overcome this problem, machine learning techniques are used to enable the recognition of previously trained objects on an image. Recently, IBM, Google and Clarifai introduced their image recognition services as cloud based services, which are investigated in this topic. Furthermore Stanford university introduced an image recognition approach, which shows recent techniques to tackle image recognition problems.</p>	
References:	
<p>https://cloud.google.com/vision/,</p>	

<http://cs.stanford.edu/people/karpathy/deepimagesent/>,
<http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/visual-recognition.html>,
<https://www.clarifai.com/>,

Applicable for BSc: yes

Applicable for MSc: yes

Further Notes: no

Topic 27: Query Languages for Graph DBs

Abstract:

The emerging of big data drives the need of network structured storage solutions, named graph databases. In order to query graph structured data, new query languages are created recently in order to enable real-time queries. Currently, SPARQL, Gremlin, and Cypher are the most frequently used query languages for graph databases. This topic discusses such query languages and states the differences to relational database query languages such as SQL.

References:

<https://www.w3.org/TR/sparql11-overview/>, <http://gremlindocs.smallette.documentup.com/>,
<http://neo4j.com/developer/cypher-query-language/>

Applicable for BSc: yes

Applicable for MSc: yes

Further Notes: no

Topic 28: Visualization of Multi-Property Graphs

Abstract:

Big data appears to emerge in a connected way, which is represented in a graph structure. Besides the connectivity property, such data carry a lot of further information. Such information is designed as properties connected to graph components. Those multi-property graphs are difficult to visualize, because of their large amount of carried content. This topic investigates solutions to deliver as much information to the user given large multi-property graphs.

References:

https://www.researchgate.net/profile/Hans-Joerg_Schulz/publication/274633015_A_Survey_of_Multi-faceted_Graph_Visualization/links/5523cb010cf2b351d9c33836.pdf

Applicable for BSc: yes

Applicable for MSc: yes

Further Notes: no

Topic 29: Storyline Generation from Social Media Data

Abstract:

Novel approaches enable the automated generation of storylines from data found on social networks, such as Twitter. The knowledge produced this way can help to make real-world processes more transparent and visible.

References:

<http://arxiv.org/pdf/1605.05195.pdf>
<http://arxiv.org/pdf/1606.03561.pdf>

Applicable for BSc: no

Applicable for MSc: yes

Further Notes: no

Topic 30: Sub-Story Detection

Abstract:

Sub-story detection allows to divide stories found on social networks into subparts and thereby enable a better understanding of the story elements and matching them to related story elements.

References:	
http://arxiv.org/pdf/1504.07361.pdf	
http://arxiv.org/pdf/1605.05894.pdf	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 31: Convolutional Neural Networks	
Abstract:	
Convolutional neural networks (CNNs) utilize layers with convolving filters that are applied to features. They have been shown to be effective for natural language processing and have achieved excellent results in semantic parsing.	
References:	
http://arxiv.org/pdf/1408.5882.pdf	
https://pdfs.semanticscholar.org/eba3/6ac75bf22edf9a1bfd33244d459c75b98305.pdf	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

4 Internet of Things & Machine-2-Machine Communication

Topic 32: Opportunistic Networking	
Abstract:	
Based on the idea of Mobile Ad-hoc Networks (MANETs), mobile devices such as smart phones are able to establish connections among each other spontaneously. This dynamic shape of collaboration reveals opportunities for a networking scheme that is based on opportunistic message forwarding. Similar to human interaction, two nodes can exchange information even if a route between them never exists.	
References:	
https://www.researchgate.net/profile/Andrea_Passarella/publication/3199770_Abstract_Opportunistic_Networking_Data_Forwarding_in_Disconnected_Mobile_Ad_hoc_Networks/links/55dc2fc908aed6a199ac7d58.pdf	
https://www.researchgate.net/profile/Chiara_Boldrini/publication/221453794_ContentPlace_social-aware_data_dissemination_in_opportunistic_networks/links/0046352458657165ff000000.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 33: Participatory Sensing Systems	
Abstract:	
The idea behind participatory sensing is to take advantage of the various sensing capacities owned by individuals or communities. Instead of deploying sensing nodes dedicated to a particular application scenario, applications may utilize the sensors already integrated into the smart phones most of us use every day. The objective of this topic is to introduce architectures for participatory sensing and discuss related issues such as security and privacy.	
References:	
http://escholarship.org/uc/item/19h777qd	
https://www.emilianodc.com/PEPSI/	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 34: Opportunistic Computing
--

Abstract:	
Based on the Opportunistic Networking paradigm, opportunistic computing allows devices to spontaneously share resources. This pattern matches the way humans interact with each other and thus provides opportunities to utilize the resources of mobile devices for human-centric computing.	
References:	
http://teaching.csse.uwa.edu.au/units/CITS4419/lectures/opportunistic-computing.pdf http://www.redpel.com/SPOC%20A%20Secure%20and%20Privacy-preserving%20Opportunistic%20Computing%20Framework%20for%20Mobile-%20Healthcare%20Emergency.pdf http://s3.amazonaws.com/academia.edu.documents/43662178/SCAMPI_service_platform_for_social_aware20160312-31562-199ivzi.pdf?AWSAccessKeyId=AKIAJ56TQRTWSMTNPEA&Expires=1474802513&Signature=YITeXfzSXElVVdmMpDjR0ZIUJA%3D&response-content-disposition=inline%3B%20filename%3DSCAMPI_service_platform_for_social_aware.pdf	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 35: VSN - Virtual Sensor Networks	
Abstract:	
Motivated by the increasing capabilities of the nodes participating in WSNs, Cloud characteristics like pooling and allocating a subset of the resources to certain tasks on-demand are applied to sensor networks. Virtual Sensor Networks (VSNs) assume that the overall set of nodes in a WSN can be virtually grouped into subsets and dedicated to specific applications or tasks. Unlike traditional WSNs, where usually all nodes perform similar tasks, a subset of nodes can be allocated for a given time period to perform specific tasks or react to the current environment. A sensor network deployed in an area recovering from a disaster may, for instance, observe specific events in a certain region. Using VSN, nodes close to that region could be grouped and further investigate the situation by executing specific tasks.	
References:	
http://www.mdpi.com/1424-8220/12/2/2175 http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4151668	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 36: Sensor Virtualization	
Abstract:	
Sensor virtualization aims at overcoming the resource constraints of traditional WSN by integrating Cloud resources and providing access to the physical sensors to multiple users. Sensor Virtualization approaches allow simplifying the management of and access to the sensors using standard Cloud interfaces and applications that hide the actual location and the diversity of the sensors from the consumers (e.g. abstraction from technical details, focus on the semantics, grouping of functional identical sensors).	
References:	
http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5678060 http://www.hindawi.com/journals/ijdsn/2013/917923/	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 37: TEDS – Transducer Electronic Data Sheets

Abstract:	
IEEE 1451 is a collection of standards that aims at providing uniform interfaces to interact with sensors or actuators (transducer) independently of the underlying communication technology. The core element of the standard collection is the definition of the Transducer Electronic Data Sheet (TEDS), which contains the information required by a measurement system to interact with a transducer (e.g. ID, calibration and correction data) and therefore allows integrating the transducer in a plug and play manner. The TEDS can be stored within the memory of the transducer or as a separate file downloadable from the internet if legacy transducers without memory need to be integrated. Compared to other device discovery and description technologies like UPnP, TEDS is more suitable for highly resource constrained devices.	
References:	
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5668466 http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4483728	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 38: Task Scheduling in MANETs	
Abstract:	
Mobile Ad-Hoc Networks describe an evolution of the well-known Sensor-Network concept. It can be assumed that some nodes participating in a MANET provide sufficient resources to establish an ad-hoc grid of computing nodes. As a result, task scheduling concepts that take the dynamics of MANETs into consideration are required.	
References:	
http://www.grc.upv.es/biblioteca/Mobile%20ad%20hoc%20networking%20imperatives%20and%20challenges.pdf http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.169.1132&rep=rep1&type=pdf	
Applicable for BSc: no	Applicable for MSc: yes
Further Notes: no	

Topic 39: Extending the Cloud: The Role of Fog and Mobile Edge Computing	
Abstract:	
Caused by the proliferation of Internet of Things (IoT) applications, a continuously increasing amount of sensors and data sources connected to information networks like the Internet can be observed. A common approach to implement such IoT applications is to collect and forward data to analytics engines usually hosted on Clouds. This generates a large amount of traffic and stresses the availability and performance of the underlying information networks. As a mitigation, recent approaches like Fog Computing or Mobile Edge Computing investigate in utilizing resources offered by edge devices like routers or smart phones. The objective is to significantly reduce the amount of network load by implementing pre-processing of data close to the sources. The aim of this topic is to give an introduction to Fog and Edge Computing including architecture approaches and discuss the approaches with regard to common distributed systems challenges like availability, fault tolerance or scalability.	
References:	
http://www.openfogconsortium.org/ - OpenFog Reference Architecture White Paper (document available on request) http://www.etsi.org/technologies-clusters/technologies/mobile-edge-computing - MEC White Paper http://s3.amazonaws.com/academia.edu.documents/46135664/SIGCOMM-MMC-Fog.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1474944127&Signature=ZgJ04wcBmlhsBX5YSTljzDh0xuQ%3D&response-content-disposition=inline%3B%20filename%3DFog_Computing_and_Its_Role_in_the_Intern.pdf	

Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

5 Miscellaneous Topics

Topic 40: Java 8 Stream API	
Abstract:	
Version 8 of the Java Platform introduced features that allow developers to leverage the advantages of a functional programming based syntax (lambda expressions). Furthermore, the new Stream API allows to process data in a declarative way and provides features to utilize multi-core architectures without having to write multithreaded code. The objective of this topic is to introduce the fundamental, functional programming inspired, concepts of the Java 8 stream API and lambda expressions to the fellow students.	
References:	
http://www.oracle.com/technetwork/articles/java/ma14-java-se-8-streams-2177646.html http://www.oracle.com/webfolder/technetwork/tutorials/obe/java/Lambda-QuickStart/index.html http://www4.fh-swf.de/media/downloads/fbtbw/download_8/devries_1/Funktionen-Streams.pdf (german)	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 41: Micro-Services	
Abstract:	
The term Micro-Services is related to a style or pattern in software architecture. Single applications are developed as a suite of small services. These services interact with each other using inter process communication and language agnostic APIs. The objective of this topic is to introduce the micro-services architecture style and distinguish it from related styles such as SOA.	
References:	
http://martinfowler.com/articles/microservices.html http://nirmata.com/2015/02/microservices-five-architectural-constraints/ http://injoit.org/index.php/j1/article/view/139	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: no	

Topic 42: Netflix Open Source Software Center	
Abstract:	
Netflix uses various Open Source Software (OSS) to serve the real-time load to over 62 million users. Major challenges are the maintenance of data persistence, build environments, security, big data analytics, reliability and performance of the service at massive scale. Netflix bundles the used software and provide the collection as Netflix Open Source Software Center.	
References:	
https://netflix.github.io/ http://techblog.netflix.com/	
Applicable for BSc: yes	Applicable for MSc: yes
Further Notes: This topic is related to the Micro-Services topic. The presentation should be prepared with regard to the presence of the more general Micro-Services presentation in the same	

seminar. Thus, if no Micro-Services talk will be given by a fellow student, a short introduction to the topic may be eligible. Otherwise, the a Micro-services introduction may be skipped.

Topic 43: Distributed Module Systems – R-OSGi

Abstract:

OSGi is a specification of a platform-independent (Java based) framework for modularized development of applications and services based on the principle of component-based software engineering [68]. A component, sometimes also referred to as a module, is a package of software that encapsulates related functions and data and provides these as services to other components using well defined interfaces. A component provides the implementation of a certain set of system processes defined by an interface, which is accessible by other components. According to the modularity of a system, an important capability of components is their exchangeability. OSGi allows updating or replacing a component at runtime without breaking the system. Thus, OSGi provides a dynamic service execution environment that is able to deploy and wire services in shape of components at runtime. The objective of this topic is to introduce concepts such as R-OSGi and distributed OSGi. These concepts aim to extend the so far centralized OSGi approach with features for distributed module management.

References:

<http://people.inf.ethz.ch/troscoe/pubs/middleware07-rosgi.pdf>
<https://cxf.apache.org/distributed-osgi.html>

Applicable for BSc: yes

Applicable for MSc: yes

Further Notes: no

Topic 44: OAuth2.0 and OpenID-Connect

Abstract:

OAuth2.0 is an authorization protocol that follows an interesting approach with regard to the peer to peer resource sharing. When using Cloud services, we often face situations where we grant access to a service to resources hosted by another service (e.g. a social network service that integrates with an e-mail service). This means the resource requesting client (i.e. the third party operating the Cloud service) differs from the resource owner (i.e. the user). Thus, the resource owner usually has to share its credentials with the third party, which introduces serious obstacles (access cannot be revoked without revoking access of all third parties, third party usually stores the user's credentials, validity of access is difficult to manage). OAuth separates the client from the resource owner and issues. OpenID-Connect provides an authentication layer on top of the OAuth2.0 protocol. The objective of this topic is to introduce the basic concepts of Oauth2.0 and OpenID-Connect with regard to operating secure Cloud services.

References:

<http://oauth.net/2/>
<http://openid.net/connect/>
<http://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText=Inside%20the%20identity%20management%20game>

Applicable for BSc: yes

Applicable for MSc: yes

Further Notes: no

Topic 45: HL7 - Continuity of Care Document

Abstract:

The Continuity of Care Document (CCD) is a joint effort of HL7 International and ASTM. CCD fosters interoperability of clinical data by allowing physicians to send electronic medical information to other providers without loss of meaning and enabling improvement of patient care. CCD is an

implementation guide for sharing Continuity of Care Record (CCR) patient summary data using the HL7 Version 3 Clinical Document Architecture (CDA). (Source: [1])

References:

[1] http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6

[2] <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.516.1037&rep=rep1&type=pdf>

[3] <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4447096/>

Applicable for BSc: yes

Applicable for MSc: yes

Further Notes: no