Big Data Compression

Big Data Compression

Modulnummer:

Bachelor

Pflicht/Wahl ☒
Wahlpflicht ☐
Wahl ☐
Sonderfall ☐

Modulbereich: Pflicht

Anzahl der SWS	V	UE	K	S	Prak.	Proj.	Σ
	0	0	4	0	0	0	4

Kreditpunkte: 6

Turnus
unregelmäßig

Formale Voraussetzungen: Keine

Inhaltliche Voraussetzungen: -

Vorgesehenes Semester: ab 1. Semester

Sprache: Deutsch

Ziele:

- Fähig sein, Kodierungen anzuwenden und zu implementieren.
- In der Lage sein, wichtige Eigenschaften von Kodierungen zu überprüfen.
- Im Stande sein, verschiedene Integer-Kodierungen anzuwenden und zu implementieren.
- In der Lage sein, die verbreitesten File Compressors (gzip, bzip, ZIP) zu verstehen und rudimentär implementieren zu können.
- Fähig sein, komprimierte Selbst-Indexe zu erstellen und effizient darauf Suche zu betreiben.
- Im Stande sein, grammatikbasierte Komprimierungsverfahren anzuwenden und zu implementieren.
- In der Lage sein, Algorithmen effizient direkt auf komprimierten Daten, ohne vorherige Dekomprimierung, laufen zu lassen.

Inhalte:

1. Shannon Entropy, k-th order Entropy
2. Optimal Prefix Codes, Huffman Codes
3. Kraft Inequality, Integer Codes
4. Arithmetic Coding
5. Dictionary Encoding (Lempel-Ziv 77, LZ78, LZW)
6. Burrows-Wheeler Transform
7. Grammar-Based Compression
8. Algorithms on Compressed Strings
9. Tree Compression (DAGs, TreeRePair)
10. Algorithms on Compressed Trees

Unterlagen (Skripte, Literatur, Programme usw.): Werden in der Veranstaltung bekanntgegeben

Form der Prüfung:
s. Prüfangebot

Arbeitsaufwand		
	Präsenz	56 h
	Übungsbetrieb/Prüfungsvorbereitung	124 h
	Summe	180 h

Lehrende:
Prof. Dr. S. Maneth

Verantwortlich:
Prof. Dr. S. Maneth