The UJ Centre for Telecommunications

THE **UJ CENTRE FOR TELECOMMUNICATIONS** AND **THE INFORMATION THEORY SOCIETY CHAPTER** OF THE **SA IEEE SECTION** CORDIALLY INVITES YOU TO **A PRESENTATION BY PROF. J. WEBER, UNIVERSITY OF DELFT, NETHERLANDS**

Date:	**03 November 2017 (Friday)**
Time:	**10:30 – 12:30**
Venue:	**Techno-Lab, University of Johannesburg**
RSVP Contact:	**Wendy Smith : wendys@uj.ac.za by 31 October 2017**

Programme:

10:30 – 11:00	**Informal Discussion (Tea, Coffee)**
11:00 – 12:00	**PRESENTATION: PROF. J. WEBER**
12:00 – 12:30	**Finger Lunch and Further Discussions**

Title

Coding for Distributed Storage: Bounds for Cooperative Locality

Abstract

In cloud storage, data is stored on multiple nodes at different locations. The effects of localized disruption of service can be effectively mitigated if the system has the ability to recover the data stored at a failed node by accessing other nodes that form a repair set for the failed node. In order to achieve this, coding is employed where data stored at a failed node is considered as an erasure in coding-theoretic terminology. The ability of a code to recover from node failures is measured by the well-known concept of Hamming distance. Inevitably coding introduces storage overhead to store redundant data and transmission overhead to exchange information between nodes in order to recover the lost data. Storage overhead is measured in terms of redundancy, a classical coding-theoretic concept. As a measure of transmission overhead, Gopalan et al. introduced the new concept of locality, which is the number of nodes that need to be accessed in the repair process. In particular, a code has r-locality if the data stored at any given node can be recovered by accessing at most r other nodes, i.e., each node has a repair set of size at most r.

The above concept of locality assumes that only one node fails. This guarantees that all nodes in a repair set of a failed node are reliable. However, based on practical considerations, it is natural to address the case of more than one failed node. One approach proposes having multiple disjoint repair sets, each of size at most r. In particular, if each node has e disjoint repair sets and if the total number of failed nodes in the system is at most e, then each failed node has at least one repair set that does not contain any failed nodes. Another approach associates to each node a set of at most $r+e-1$ other nodes such that if the node fails and up to $e-1$ nodes in the set also fail, the remaining nodes in the set form a repair set for the node associated with the set. A third approach assigns to each set of e nodes a disjoint repair set of at most r nodes, called cooperative repair set. If up to e nodes fail, then a repair set associated with the failing nodes can be used to recover the data stored at the e-failed nodes. In this approach, the e-failed nodes are not repaired independently as in the previous two approaches where r nodes are involved in the repair process of each failing node, but rather collectively as r nodes are involved in the repair process of all the failed nodes. Therefore, a code achieving this requirement is said to have (r,e)-cooperative locality. For each one of the three approaches, considerable literature is devoted to both the study of bounds on the code's length, redundancy, Hamming distance, locality, and the number of failed nodes allowed, as well as the construction of codes that are optimal in the sense of achieving these bounds. In case $e=1$, these three approaches reduce to the concept of r-locality as proposed by Gopalan et al.

In this presentation, a general introduction is given on the above-described locality concepts and, furthermore, bounds on linear codes with (r,e)-cooperative locality are presented, based on joint work by Khaled Abdel-Ghaffar and the presenter. Our approach is based on the concept of generalized Hamming weights, proposed by Wei as a generalization of Hamming distance of a linear code. This allows us to generalize many known bounds on such codes leading to bounds which, for some code parameters, improve upon the tightest bounds known in the literature.

Biography: J Weber

Jos H. Weber was born in Schiedam, The Netherlands, in 1961. He received the M.Sc. (in mathematics, with honors), Ph.D., and MBT (Master of Business Telecommunications) degrees from Delft University of Technology, Delft, The Netherlands, in 1985, 1989, and 1996, respectively. He is a senior member of IEEE.

Since 1985, he has been with the Delft University of Technology. Currently, he is an associate professor at the Department of Applied Mathematics. He is the chairperson of the Werkgemeenschap voor Informatie- en Communicatietheorie since 2006, and the secretary of the IEEE Benelux Chapter on Information Theory since 2008. He was a Visiting Researcher at the University of California (Davis, CA, USA), the Tokyo Institute of Technology (Japan), the University of Johannesburg (South Africa), and EPFL (Switzerland). His main research interests are in the area of channel coding. More information is available at http://homepage.tudelft.nl/2k00k/.