§44. Analysis of MHD Stability Using Data Mining Technique in LHD Plasmas

Yamamoto, S. (Institute of Advanced Energy, Kyoto Univ.), Blackwell, B., Pretty, D. (Australian National Univ.), Ohdachi, S., Sakakibara, S., Watanabe, K.Y., Toi, K.

Data mining technique based on statistics, pattern recognition, artificial intelligence and information technology have been used in the areas of distribution and finance for business, and bio-informatics, astronomy and geology for science. Data mining techniques can extract new information because they are able to automatically pick out patterns in large amounts of high-dimensional data. We applied a data mining technique to analyze the fluctuation signals within a large database in order to identify MHD instabilities on LHD plasmas. Moreover, the entry of information about MHD instability classifications into a database enables us to exactly and quickly investigates the characteristics of MHD stability through parameter studies. We successfully classified and identified the MHD instabilities in H-1 and Heliotron J plasmas by the data mining technique.

After successful trials in late 2009, progress in the LHD data mining project of this year includes, (1) Upgrade of the software to the new version, (2) Analysis of all shots using the older 14bit digitizers (over 50,000 shots) to check uncertainties in the time base divisor, (3) Following a suggestion from the 7th CWGM, application of the technique to all ~300 LHD shots in the IEA stellarator working group high performance database DB07_25, and (4) Processing of 10000 longer shots (4 M samples/channel). The new version is a complete rewrite to allow much more flexible selection of data sources, a comprehensive *.ini style configuration file based on the RFC822 standard, a history feature to track processing steps, and simplified usage, and more complete documentation. An improvement in speed of about 10x can be obtained by writing data to intermediate text files instead of an SQL database. The simplicity and power of use is illustrated by the length of the script used to compile the data for checking the time base divisor in step 2 above – one line of 140 characters total. The amplitude of Mirnov signals is difficult to relate to the amplitude internal to the plasma, without knowing the free/fixed boundary solution for each mode. The new code has an estimate of amplitude which attempts to overcome the effects of normalization and variable attenuation of probes are different distances from the plasma. The IEA stellarator working group high performance database DB07_25 is the subject of several papers, and was suggested as a good starting point for application to LHD physics. The shots are in the range 27153-54198, and include some of the highest beta shots up to that shot. Toroidal mode analysis shows toroidal mode number n=0, 1 and 2 modes. Figure 2 shows that many very rapid down-sweeping chirps in this data set are captured. Although not all are captured, those that are captured are representative. Work is continuing to improve the captured fraction. Figure 3 shows all n=1 modes from the whole ~300 shot data set with the amplitude nature of the mode indicated by dot size and the plasma beta coded by color. The highest beta shots are characterized by early n=1 activity decreasing rapidly in frequency. The incorporation of time-dependent density was achieved in the visit of Dr. Blackwell to NIFS in late 2010, and this will allow determination of whether this decrease is due to Alfvénic scaling or otherwise. Further processing of the 10,000 shots of point 4 above will be performed once the time-dependent density is in place.

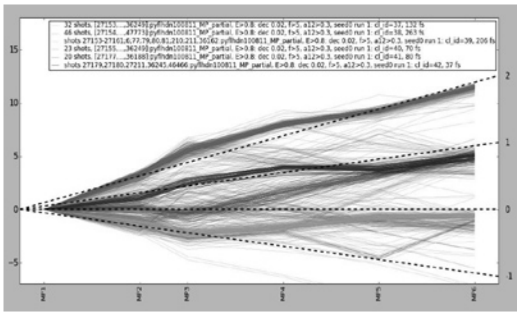

Fig. 1. Toroidal mode number analysis of DB07_25 (292 shots). Clusters of n=0,1 and 2 are evident, although the n-1 cluster is split.

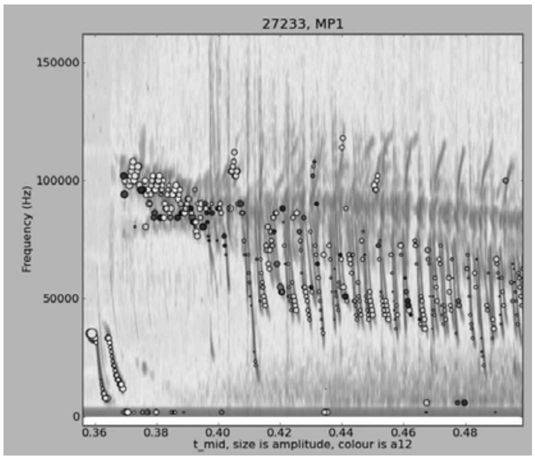

Fig. 2. Good success rate of capturing with rapid frequency down chirping.

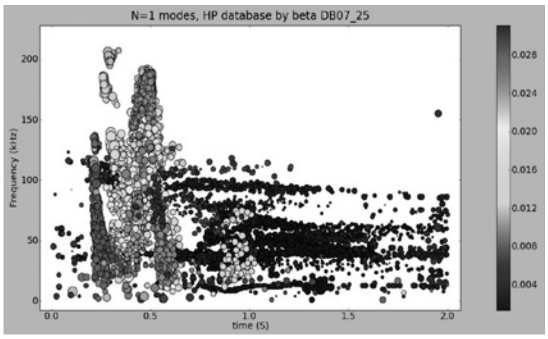

Fig. 3. n=1 modes for all shots in the high performance database..