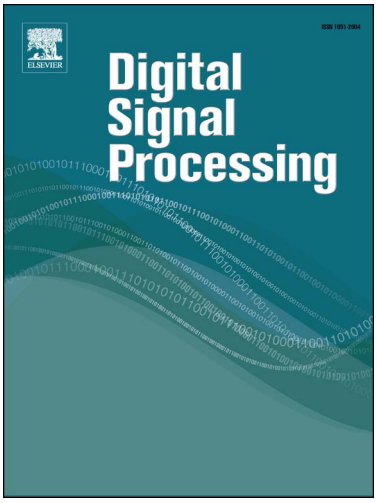
Accepted Manuscript



Impact of a priori MS/MS intensity distributions on database search for peptide identification

Hatem Loukil, Mohamed Tmar, Mahdi Louati, Afif Masmoudi, Faiez Gargouri

PII: S1051-2004(17)30081-7
DOI: http://dx.doi.org/10.1016/j.dsp.2017.04.012
Reference: YDSPR 2108

To appear in: *Digital Signal Processing*

Please cite this article in press as: H. Loukil et al., Impact of a priori MS/MS intensity distributions on database search for peptide identification, *Digit. Signal Process.* (2017), http://dx.doi.org/10.1016/j.dsp.2017.04.012

Impact of a priori MS/MS intensity distributions on database search for peptide identification

Hatem Loukil[a,*], Mohamed Tmar[a], Mahdi Louati[b], Afif Masmoudi[c], Faiez Gargouri[a]

[a] *MIRACL, Multimedia, InfoRmation systems and Advanced Computing Laboratory, University of Sfax, CP 3021, Tunisia*
[b] *National School of Electronics and Telecommunications, Sfax Technopark, BP 1163, CP 3018, Tunisia*
[c] *Laboratory of Probability and Statistics. Faculty of Sciences, University of Sfax, BP 1171,CP 3038, Tunisia*

Abstract

Many database search methods have been developed for peptide identification throughout a large peptide data set. Most of these approaches attempt to build a decision function that allows the identification of an experimental spectrum. This function is built either starting from similarity measures for the database peptides to identify the most similar one to a given spectrum, or by applying useful learning techniques considering the database itself as a training data. In this paper, we propose a peptide identification method based on a similarity measure for peptide-spectrum matches. Our method takes into account peak intensity distribution and applies it in a probabilistic scoring model to rank peptide matches. The main goal of our approach is to highlight the relationship between peak intensities and peptide cleavage positions on the one hand and to show its impact on peptide identification on the other hand. To evaluate our method, a set of experiments have been undertaken into two high mass spectrum accuracy data sets. The obtained results show the effectiveness of our proposed approach.

Keywords: Mass spectrometry, Tandem mass spectrum, Peptide

*Corresponding author.
Email addresses: loukil.hatem@hotmail.com (Hatem Loukil), mohamedtmar@isimsf.rnu.tn (Mohamed Tmar)