A NETWORK INTRUSION DETECTION METHOD USING DEMPSTER-SHAFER THEORY

Ladislav Beránek, Jiří Knížek***

*University of South Bohemia in Ceske Budejovice

**Charles University in Prague, Faculty of Medicine in Hradec Kralove

Abstract:

An intrusion detection system (IDS) detects unauthorized manipulations of computer systems. Operation as feature reduction (including feature extraction and feature selection) plays an important role in the sense of improving classification performance and reducing the computational complexity of intrusion detection system. Feature reduction is even more important at online detection when less computational power and fast real time delivery compared with offline detection is needed. In this paper, Dempster Shafer theory based on KNN analysis approach [2] is applied to feature extraction in online network intrusion detection problem. We used the KDD Cup 99 [1] data and reduced its 41 features such that significant less number of features would be fed into this classifier.

We used theoretical approach of Denoeux [2] which combines in his work Dempster Shafer theory [3] evidence coming from the k nearest neighbors of a test example (part of our data). Denoeux also addresses ambiguity and distance rejection, and uncertainty and imprecision in class labels [3]. We used this classifier and studied the performance of the Dempster-Shafer theory based KNN classifier used for intrusion detection system (KDD Cup 99 data).

The whole paper is organized as follows: section 2 briefly describes the Dempster-Shafer theory of evidence and the KNN classifier based on this theory. In the next section, the data analysis methods including preprocessing and feature extraction techniques are described. In section 4, the voting KNN classifier, distance-weighted KNN classifier, and Dempster-Shafer KNN classifier are compared on an (KDD Cup 99 [1] data) and the conclusion is given in this paper.

It has been shown that the Dempster-Shafer KNN classifier will result in higher classification accuracy in comparison with other two KNN classifiers. Thanks to its simplicity and performance, we are now looking forward to evaluating the performance of the real intrusion detection system (IDS) implemented on our university.

We use MATLAB software for computation and visual exploration of data. This software is very suitable for these analyses considering its capability of data management and transformation tools ranging from graph procedures to a full-featured matrix algebra language. We present also the detail algorithms in MATLAB.

Reference:

[1] Kdd cup. the third international knowledge discovery and data mining tools competition. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[2] T. Denoeux, “A k-nearest neighbor classification rule based on Dempster-Shafer theory,” IEEE Trans. Syst. Man Cybern., vol. 25, no. 5, pp. 804–813, 1995.

[3] G. Shafer, A mathematical theory of evidence. Princeton university press Princeton, NJ, 1976.

Contact information:
Department of Applied Mathematics and Informatics, Faculty of Economics, University of South Bohemia in Ceske Budejovice, Studentska 13, 370 05 Ceske Budejovice, Czech Republic, e-mail: beranek@ef.jcu.cz