features a_1, a_2, and a_5 can contribute more information than the sum of each features. Besides, for the real-world applications of high-throughput microarray data, finding interacting features is to search a small number of pertinent genes from hundreds of thousands of ones that reflecting the immune response mechanism involving both antigen presentation and immunoproteasome pathways [24].

For streaming feature selection, we cannot require the domain knowledge in advance and do not know the information of the entire feature space before learning. Most of the existing streaming feature selection methods, such as OSFS and SAOLA mentioned earlier, treat features individually and do not consider the interaction among features. Besides, many effective and efficient learning algorithms assume the independence of features. However, they may fail badly when the degree of feature interaction becomes critical [23].

Motivated by this, we focus on the problem of feature interaction and propose a new streaming feature selection method that can select features to interact with each other, named Streaming Feature Selection considering Feature Interaction (SFS-FI). The main contributions of this article are as follows.

1) We give the formal definition of feature interaction. Meanwhile, we analyze and demonstrate the relationship between feature relevance and feature interaction.
2) We provide a systematic analysis of two-way interaction, three-way interaction, and four-way interaction in feature selection. Meanwhile, we demonstrate that four-way interaction can be converted into the sum of some three-

Filter methods evaluate the feature importance according to certain criteria that are independent of any learning algorithms. For example, by using the concept of distance correlation, Kundu and Mitra [28] developed a novel similarity-based feature selection algorithm that does not need an exhaustive traversal of the search space. Yang *et al.* [29] investigated the incremental perspective for fuzzy rough set-based feature selection, which assumes that data can be presented in the sample subsets one after another. Wrapper methods evaluate the quality of the selected features with a predefined learning algorithm. The article [30] describes a novel wrapper feature selection algorithm for classification problems, which utilizes a genetic algorithm to wrap extreme learning machine to search for the optimum subsets in the huge feature space. Embedded methods perform feature selection in the process of model construction. The representative methods of embedded mode are regularized regression-based feature selection algorithms. Pang *et al.* [31] proposed a novel framework to solve the original $l_{2,0}$-norm constrained sparse regression-based feature selection problem. Besides, Bing *et al.* [32] present a comprehensive survey of the state-of-the-art work on evolutionary computation for feature selection and identified the contributions of these different algorithms.

For traditional feature selection methods, most of them concentrate on removing irrelevant and redundant features or selecting the most relevant features from a given candidate feature space. Among them, there are few works focusing on the feature interaction. More specifically, Jakulin and Bratko [33] focused on the assumption of inde-