formed enough to reject unseen classes of documents due to its significant open space risk.

Traditional multiclass learners optimize only on the known classes under the closed world assumption, while a potential learner for open classification has to optimize for both the known classes and for the unknown classes. Some recent research in the field of computer vision studied the problem, which they call *open set recognition* (Scheirer et al., 2013; 2014; Jain et al., 2014) for facial recognition. Classic learners define and optimize over empirical risk, which is measured on the training data. For open classification, it is crucial to consider how to extend the model to capture the risk of the unknown by preventing overgeneralization or overspecialization. In order to tackle this problem, Scheirer et al. (2013) introduced the concept of *open space risk* and formulated an extension of existing one-class and binary SVMs to address the open classification problem. However, as we will see in section 3, their proposed method is weak as the positively labeled open space is still an infinite area.

In this work, we propose a solution to reduce the open space risk while also balancing the empirical risk for open classification. Intuitively, given a positive class of documents, our open space for the positive class is considered as the space that is sufficiently far from the center of the positive documents. In the multiclass classification setting, each of the *m* target classes is surrounded by a ball covering the positively labeled (the target class) area, while any document falling outside of all the *m* balls is considered belonging to the unknown class.

Recent work by Fei and Liu (2015) proposed a new learning strategy called *center-based similarity space learning* (*CBS learning*) to deal with the problem of covariate shift in binary classification. We found that it is also suitable for open classification. Instead of conducting learning in the traditional document space (or *D-space*) with n-gram features, CBS learning learns in a similarity space. Unlike SVM learning in D-space that bounds the positive class only by an infinite half-space formed with the decision hyperplane, which has a huge open space risk, CBS learning finds a closed boundary for the positive class covering only a finite area, which is a spherical area in the original D-space and thus reduce the open space risk significantly. While discussing CBS learning, we will also describe the underlying assumptions made by it which were not stated in our earlier paper (Fei and Liu, 2015). Our final multiclass classifier is called *cbsSVM* (based on SVM).

To the best of our knowledge, this is the first attempt to study multiclass open classification in text from the open space risk management perspective. Our experiments show that *cbsSVM* for multiclass open classification produces superior classifiers to existing state-of-the art methods.

2 Related Work

Compared to research on multiclass classification with the closed world assumption, there is relatively less work on open classification. In this section, we review related work on one-class classification, SVM decision score calibration, and others.

One-class classifiers, which only rely on positive training data, are natural starting solutions to the multiclass open classification task. One-class SVM (Scholkopf et al., 2001) and SVDD (Tax and Duin, 2004) are two representative one-class classifiers. One-class SVM treats the origin in the feature space as the only member of the negative class, and maximizes the margin with respect to it. SVDD tries to place a hypersphere with the minimum radius around almost all the positive training points. It has been shown that the use of Gaussian kernel makes SVDD and One-class SVM equivalent, and the results reported in (Khan and Madden, 2014) demonstrate that SVDD and One-class SVM are comparable when the Gaussian kernel is applied. However, as no negative training data is used, one-class classifiers have trouble producing good separations. We will see in Section 4 that their results are poor.

This work is also related to using thresholded probabilities for rejection. As the decision score produced by SVM is not a probability distribution, several techniques have been proposed to convert a raw decision score to a calibrated probability output (Platt, 2000; Zadrozny and Elkan, 2002; Duan and Keerthi, 2005; Huang et al., 2006; Bravo et al., 2008). Usually a parametric distribution is assumed for the underlying distribution, and raw scores are mapped based on the learned model. A variation of Platt's (2000) approach is the most widely used probability estimator for SVM score calibration. It fits a sigmoid function to the SVM scores during training. Provided with a threshold, a test instance can be rejected if the highest probabil-