

распространенным алгоритмом построения тематических моделей является LDA, лишенный практически всех недостатков своих предшественников. Латентное размещение Дирихле (LDA, Latent Dirichlet Allocation) – это модель, объясняющая результаты наблюдений с помощью неявных групп, что позволяет получить объяснение, почему некоторые части данных схожи. Например, если наблюдениями являются слова, собранные в тексты, утверждается, что каждый текст представляет собой смесь небольшого количества тем и что появление каждого слова связано с одной из тем документа. LDA впервые был представлен в качестве графической модели для обнаружения тем Дэвидом Блеем, Эндрю Нг и Майклом Джорданом в 2002 году. Модель LDA устраняет такие недостатки популярных ранее моделей, как склонность к переобучению, невозможность вычислить вероятность нового документа и отсутствие закономерности при генерации документов из сочетания полученных тем.

Литература

1. Blei D., Ng A., Jordan M. Latent Dirichlet Allocation // Journal of Machine Learning Research. – MIT Press, 2003. – No. 3. – P. 993–1002.
2. Воронцов К. В. Математические методы обучения по прецедентам (теория обучения машин) // Курс лекций ВМК МГУ и МФТИ. – 2011.

ИСПОЛЬЗОВАНИЕ ВЕРОЯТНОСТНОГО ЛАТЕНТНО-СЕМАНТИЧЕСКОГО АНАЛИЗА ДЛЯ ПОСТРОЕНИЯ ВЕРОЯТНОСТНЫХ ТЕМАТИЧЕСКИХ МОДЕЛЕЙ ТЕКСТОВЫХ КОЛЛЕКЦИЙ

М.И. Селюк, Н.Н. Шинкевич, М.В. Стержанов

Одной из важных задач автоматической обработки текстов является кластеризация текстовых коллекций. Для выявления скрытых тем в текстовых коллекциях в последнее время все чаще применяются вероятностные тематические модели. Одним из наиболее часто используемых алгоритмов для построения тематических моделей является PLSA.

Вероятностный латентно-семантический анализ (индексирование) (PLSA, Probabilistic Latent Semantic Analysis) был предложен Томасом Хоффманом в 1999 году. В основе метода лежит аспектная модель (aspect model), которая связывает скрытые (латентные) переменные тем с каждой наблюдаемой переменной словом или темой. Задача состоит в выявлении латентных переменных. Таким образом, каждый документ может относиться к некоторым темам с некоторой вероятностью, что является отличительной особенностью этой модели по сравнению с подходами, не основанными на вероятностном моделировании.

Достоинством метода можно считать его способность выявлять зависимости между словами, когда обычные статистические методы бессильны. PLSA также может быть применен как с обучением (с предварительной тематической классификацией документов), так и без обучения, что зависит от решаемой задачи. К недостаткам модели можно отнести склонность к переобучению и неприменимость к большим наборам данных, невозможность вычислить вероятность документа, которого нет в наборе данных, а также отсутствие какой-либо закономерности при генерации документов из сочетания полученных тем.

Литература

1. Manning C. D. Introduction to Information Retrieval // MIT Press, 2008.
2. Hofmann Thomas. Probabilistic latent semantic indexing // In Proc. of the SIGIR99.

ЭЛЕКТРОННЫЕ СВОЙСТВА ДЕФЕТНЫХ СТРУКТУР ФОСФОРЕНА

В.А. Скачкова, М.С. Баранова, Д.Ч. Гвоздовский

Точечные дефекты и примеси часто оказывают значительное влияние на физические свойства материала. Экспериментальное определение дефектов обычно крайне сложное и косвенное, кроме того, требует невероятных сочетаний различных методов. При исследовании дефектов, первопринципные методы показали себя как мощный теоретический подход, достаточно надежный, чтобы использоваться как предсказательный инструмент. Проведено квантово-механическое моделирование влияния моновакансии в слое черного фосфора (фосфорене), на его

структурные свойства. Расчеты проводились в программном комплексе VASP, который реализует метод теории функционала электронной плотности [1]. Данные моделирования показали, что наличие вакансии приводит к образованию трех оборванных связей, две из которых, благодаря гибкости фосфорена, соединяются друг с другом, а третья может иметь положительный, отрицательный и нейтральный заряд. Энергия образования нейтральной вакансии составила 1,88 эВ. Расчет электронной структуры нейтральной вакансии показал, что фосфорен теряет свою прямозонную природу за счет появления разрешенных состояний в запрещенной зоне.

Литература

1. Kresse, G. VASP the guide: tutorial / G. Kresse. –Vienna, 2003.

МНОГОСЛОЙНЫЕ ПЛЕНОЧНЫЕ ЭКРАНЫ НА ОСНОВЕ СПЛАВОВ NiFe ДЛЯ ЗАЩИТЫ ОТ ВНЕШНИХ МАГНИТНЫХ ПОЛЕЙ

А.А. Солобай, С.С. Грабчиков, А.В. Труханов

Над решением проблемы защиты изделий электронной и радиоэлектронной техники от воздействия различного типа внешних электромагнитных полей работают ведущие специалисты многих стран мира [1–3]. Известно, что наиболее эффективным способом защиты является экранирование. Была исследована эффективность экранирования переменных электромагнитных и магнитостатических полей многослойными пленочными структурами на основе сплавов NiFe, полученных методом электролитического осаждения, с целью разработки высокоэффективной защиты корпусов приборов и аппаратуры широкого спектра назначения. Было отмечено, что с ростом толщины экранов эффективность возрастает, что связано с увеличением площади экрана и, соответственно, снижением магнитного сопротивления, а пик максимума эффективности смещается в область более высоких магнитных полей. Для сопоставления полученных результатов с характеристиками промышленных материалов были измерены полевые зависимости экранов, изготовленных из промышленной аморфной ленты 84KXCP и электролитического сплава Ni₈₀Fe₂₀. Обнаружено, что магнитная проницаемость аморфной ленты высока только в ограниченном диапазоне, а затем резко падает, в отличие от магнитных электролитически осажденных пленок, которые эффективны в более высоких полях, начиная с 200 А/м. Сделан вывод, что при подборе изготовлении экранов для магнитостатической защиты следует учитывать как основные магнитные характеристики, так и роль неоднородности магнитного поля в объеме материала экрана, и связанную с этим нелинейность магнитной проницаемости.

Литература

1. Апполонский, С.М. / Расчеты электромагнитных полей / С.М. Апполонский, А.Н. Горский. – М.: Маршрут, 2006. – 992 с.
2. Gabower J.F. Electromagnetic interference shields for electronic devices: pat. USA № 7358447.
3. Кечиев, Л.Н. Экранирование технических средств и экранирующие системы / Л.Н. Кечиев, Б.Б. Акбашев, П.В. Степанов. – М.: Группа ИТД, 2010. – 470 с.

ПРОЕКТИРОВАНИЕ И ДИЗАЙН НА ОСНОВЕ ЦИФРОВЫХ ПРОТОТИПОВ AUTODESK ALIAS

В.А. Столер

Известно, что успех любого изделия определяется его качеством, стоимостью и брендом. Вместе с тем, в настоящее время тенденции таковы, что, более важную роль играет скорость создания модели изделия и его дизайна – они являются главным преимуществом продукта. В докладе описываются особенности создания моделей различных изделий, в том числе защитных экранов ЭМИ, используя их цифровые прототипы и дизайн, разработанные на основе программы Autodesk Alias – популярной программы компании Autodesk (США).

Так как моделирование и дизайн определяют сколько времени и ресурсов придется потратить для создания изделия, то применение Alias делает ее незаменимой графической программой с возможностью получения математических моделей. Прогрессивная технология