4. LDA позволяет определить для документа не одну тему, а несколько, чего лишен метод k-means.

Таким образом, можно сделать вывод, что для задачи вероятностного тематического моделирования лучше использовать вероятностные алгоритмы, а именно для данной задачи имеет место методы LDA (или же его собрат PLSA). При этом в результате работы, обучения на корпусе документов, полученные распределения (скрытые переменные) можно использовать для эффектной визуализации данных: построения облака слов для каждой темы.

Список использованных источников:

1. Луис Педро Коэльо, Вилли Ричард - Построение систем машинного обучения на языке Python,
2. Christopher D. Manning, Hinrich Schütze - Foundations of Statistical Natural Language Processing,
3. Diane J. Hu - Latent Dirichlet Allocation for Text, Images, and Music.

ДЕМОНСТРАЦИЯ ПРИНЦИПА РАБОТЫ ШИФРОВАЛЬНОЙ МАШИНЫ «ЭНИГМА»

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Коршунов А.А.

Стройникова Е.Д. – ассистент кафедры информатики

Для демонстрации принципа работы шифровальной машины «Энигма» разработана учебная программа с использованием языка программирования C# на основе алгоритма шифрования.

Со времён изобретения письменности людьми движет желание скрыть написанное от посторонних глаз. Было изобретено много простых способов защитить текст, адресованный конкретному человеку, таких, как стеганография и простейшие шифры перестановки («атбаш»).

Как показала практика, стеганографию не всегда возможно применить, а простейшие шифры очень быстро поддаются криптоанализу. Вследствие этого криптографы стали придумывать различные занимательные способы запутать своих «противников». Появились такие шифры, как аффинный, полиалфавитный (простейший – диск Альберти). Данные способы сокрытия информации были намного более эффективными, т.к. давали большее количество комбинаций, а соответственно усложняли расшифровку. В механизме шифровальной машины «Энигма» используется некое подобие диска Альберти, объединённого с шифратором Джефферсона.

В стандартной механической версии данной машины использовано 3 ротора и 1 рефлектор. На каждом из роторов имеется 2 алфавита: «принимающий» букву и «отдающий» букву. Соответственно на принимающем роторе имеем обычный алфавит, на отдающем – шифроалфавит. Буква, попадающая в данный механизм, проходит каждый из роторов в одном направлении, доходит до рефлектора и возвращается обратно, после чего происходит поворот заданного (не обязательно первого) ротора на 1 позицию. Таким образом, с помощью несложных преобразований, ускоренных электрической схемой, получаем несколько раз зашифрованную букву. Следует заметить, что после каждого поворота ротора шифроалфавит будет изменён. Следовательно, можно не бояться, что одна и та же буква открытого текста будет зашифрована одинаково более 2-х раз подряд. Также данная машина даёт возможность изменить входящие в неё буквы ещё до преобразования. Это как бы добавляет ещё один ротор, но уже статический. Конечно же, стоит отметить, что все роторы «Энигмы» можно ставить на любую позицию, что позволяет варьировать начало шифрования. На рисунке 1 приведена иллюстрация зашифровывания буквы в более ранних версиях «Энигмы»:

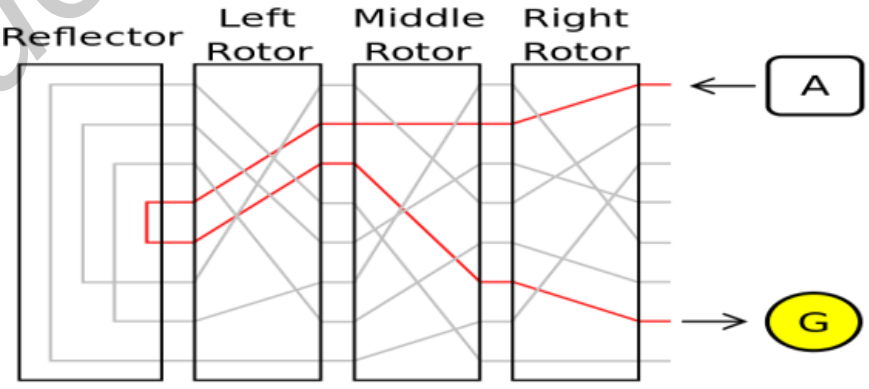

Рис. 1 – Процесс шифрования

Для воссоздания принципа шифрования машины «Энигма» был использован язык программирования C# из-за его простоты и удобства в работе с объектами. Для большей наглядности процесса для входящего и выходящего текста использовались массивы символов и из метода возвращалась буква шифроалфавита, индекс которой соответствовал бы индексу входящей буквы. В предлагаемой программной версии опущены математические преобразования, проводимые с буквой при переходе с ротора на ротор, которые присутствовали на более поздних «Энигмах». Также в программе для упрощения восприятия опущено разбиение шифротекста на блоки типа «XXXXXXXXXXXX». Результат работы программы при попытке зашифровать фразу «attackatdawn» приведён на рисунке 2:

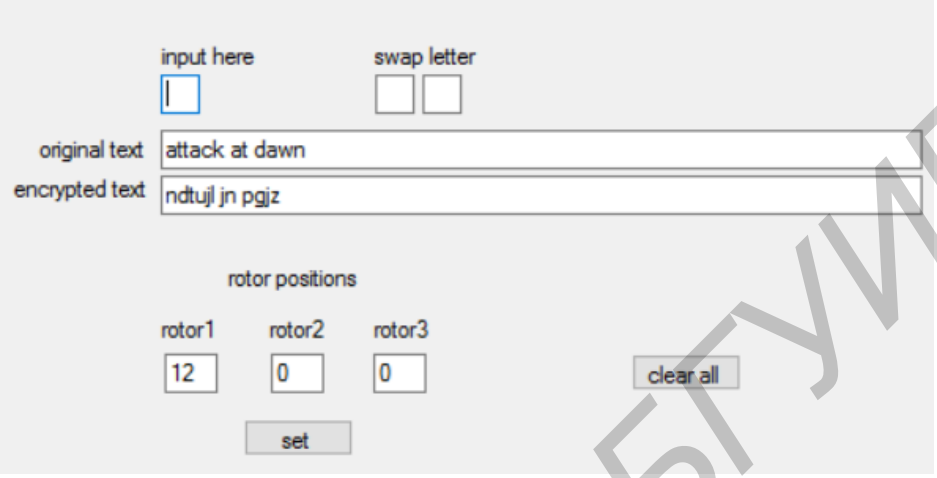

Рис. 2 – Результат работы программы

Основные преимущества данного шифра:
- простота реализации;
- большое количество комбинаций символов шифротекста;
- шифр не поддаётся сугубо математическому анализу (на данный момент).

Основной недостаток данного шифра: если шифротекст записан на грамматически строгом языке и взломщик владеет открытым текстом – расшифровать другое сообщение, зашифрованное с той же расстановкой шифрующих элементов, представляется возможным в короткие сроки.

Данная работа будет неплохим подспорьем в преподавательской деятельности, а также для наглядной демонстрации всем желающим.

В завершение хотелось бы отметить, что некоторые тексты, зашифрованные «Энигмой» и повреждённые (оставшиеся только в зашифрованном варианте) по сей день не поддаются расшифровке и восстановлению, из-за чего мы до сих пор не знаем всего о некоторых интересных событиях Второй Мировой войны.

Список использованных источников:

1. Сингх, С. Книга шифров: тайная история шифров и их расшифровки / С. Сингх ; пер. с англ. – М. : АСТ: Астрель, 2007. – 447 с.
2. Молдовян, Н. А. Криптография: от примитивов к синтезу алгоритмов / Н. А. Молдовян, А. А. Молдовян, М. А Еремеев. – СПб. : БХВ-Петербург, 2004. — 448 с.
3. Энигма. Материал из Википедии – свободной энциклопедии [Электронный ресурс]. – 2016. – Режим доступа : https://ru.wikipedia.org/wiki/%D0%AD%D0%BD%D0%B8%D0%B3%D0%BC%D0%B0.
4. The Late Tony Sale's Codes and Ciphers Website [Электронныйресурс]. – 2004. Режим доступа : http://www.codesandciphers.org.uk/enigma/rotorspec.htm.

ОРГАНИЗАЦИЯ АВТОМАТИЗИРОВАННОГО ТЕСТИРОВАНИЯ WPF ПРИЛОЖЕНИЙ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Костюкевич В.В.

Калугина М.А.- канд. физ.-мат. наук, доцент

Рассматривается задача организации автоматизированного тестирования WPF-приложений в соответствии с наиболее популярными решениями этой же задачи в рамках веб-приложений.