



Рисунок 2 – Интерфейс работы приложения

Верхние 2 кнопки осуществляют контроль за мониторингом. Компонент Chronometer осуществляет счёт времени. В таблице указываются координаты, полученные посредством мониторинга, а фактически маршрут рисуется на основании заполненной таблицы.

Кнопка «Построить маршрут» позволяет построить маршрут.

ЛИТЕРАТУРА

1. GPS [Электронный ресурс]. – 2017. – Режим доступа: <https://ru.wikipedia.org/wiki/GPS>. Дата доступа: 18.04.2017.
2. ГЛОНАСС [Электронный ресурс]. – 2017. – Режим доступа: <https://ru.wikipedia.org/wiki/ГЛОНАСС>. Дата доступа: 18.04.2017.
3. Геолокация и картография в Android приложении [Электронный ресурс]. – 2017. – Режим доступа: <http://src-code.net/geolokaciya-i-kartografiya-v-android-prilozhenii/>. Дата доступа: 19.04.2017.

УДК 004.421.6

Студ. В.Е. Журавлёв

Науч. рук. ст.преп. Ю.О. Герман

(кафедра информационных систем и технологий, БГТУ)

ANDROID-ПРИЛОЖЕНИЕ ДЛЯ ПЕРЕВОДА ТЕКСТА НА ИНОСТРАННЫЙ ЯЗЫК

В настоящее время мобильные устройства находят свое применение практически во всех областях жизнедеятельности человека. Всё больше и больше сервисов и услуг переходят с настольных версий на веб или мобильные платформы.

На сегодняшний день людям приходится читать много литературы на иностранном языке. Чтобы скоротать время, например, в транспорте, большинство читает материал на своих мобильных телефонах. Однако незнание иностранного языка приводит к тому, что

необходимо постоянно прерываться, заходить в словарь, вводить нужное слово, что неудобно для пользователей. С целью решения подобных проблем разрабатывается мобильное приложение Copy Translate, показывающее перевод слова сразу, после того как оно скопировано в буфер. Суть использования Copy Translate заключается в следующем: Android-приложение работает постоянно в фоне как служба, следит за буфером обмена и, как только в буфер поступает новый текст, переводит его с выдачей результата.

Прототип приложения был создан в программе Adobe Experience Design [1]. Дизайн приложения сделан в стиле Google Material Design. Material Design — дизайн программного обеспечения и приложений операционной системы Android от компании Google. Идея дизайна заключается в приложениях, которые открываются и сворачиваются как карточки, используя эффекты теней. У приложений не должно быть острых углов, карточки должны переключаться между собой плавно и практически незаметно [2]. Основываясь на этом, были созданы основные блоки приложения.

Согласно Google Material, основные логические блоки следует вынести в «плитки», плитка – это прямоугольник с закругленными углами, который, используя тень, находится как бы над другими элементами управления.

Поиск, доступ в меню и включение-выключение службы вынесен в toolbar, который, опять же собран по советам Material Design Guidelines.

Программная реализация написана на языке Java на основе построенной UML-модели. Java-код создаёт сам Enterprise Architect.

В приложении имеется возможность просматривать историю переводов. Здесь же реализован поиск нужного перевода. Было решено сделать поиск нечётким. Нечеткий поиск является крайне полезной функцией любой поисковой системы. Вместе с тем, его эффективная реализация намного сложнее, чем реализация простого поиска по точному совпадению.

Нечеткий поиск - это поиск информации, при котором выполняется сопоставление искомого слова заданному образцу или близкому к этому образцу значению. Алгоритмы нечеткого поиска используются в большинстве современных поисковых систем, например, для проверки орфографии. Классическая задача нечеткого поиска может принимать следующий вид: имеется текстовая информация определенного размера. Пусть пользователь вводит слово или фразу для поиска, необходимо найти в тексте все совпадения с

заданным словом с учетом возможных допустимых различий. Например, при запросе пользователя «интернет» нужно найти «интерн», «интернат» и другие слова.

Для того, чтобы оценить сходство двух слов, в тексте используются специальные метрики нечеткого поиска. Метрикой нечеткого поиска называют функцию расстояния между двумя словами, позволяющую оценить степень их сходства в данном контексте [3].

В качестве метрик используют расстояния Хемминга, Левенштейна, Дамерау-Левенштейна. Расстояние Хемминга - это число позиций, в которых соответствующие символы двух слов одинаковой длины различны. У расстояния Хемминга есть один существенный недостаток - сравнивать можно только слова одинаковой длины. Из-за этого данную метрику практически не применяют на практике.

Расстояние Левенштейна или расстояние редактирования - это минимальное количество операций (вставки, удаления одного символа и замены одного символа на другой), необходимых для преобразования одной строки в другую. Данная метрика имеет следующие недостатки:

– при перестановке слов в предложении расстояние принимает большое значение;

– расстояние Левенштейна значительно зависит от длины слова.

Расстояние Дамерау-Левенштейна - это мера разницы двух строк символов, определяемая как минимальное количество операций вставки, удаления, замены и перестановки соседних символов, необходимых для перевода одной строки в другую. Данная метрика отличается от расстояния Левенштейна только добавлением новой операции (перестановки). Расстояние Дамерау — Левенштейна $d_{a,b}(|a|, |b|)$ между двумя строками a и b определяется функцией вида [3]:

$$d_{a,b} = \min(X_x + 1, X_y + 1, X_z + C_{\text{замены}}, X_m + C_{\text{транспозиции}})$$

$$C_{\text{замены}} = \begin{cases} 1, \text{ если } S_1[i] \neq S_2[j] \\ 0, \text{ иначе} \end{cases}, \quad (1)$$

$$C_{\text{транспозиции}} = \begin{cases} 1, \text{ если } S_1[i] = S_2[j-1] \text{ и } S_1[i-1] = S_2[j] \\ \infty, \text{ иначе} \end{cases}$$

где X_x , X_y , X_z , X_m – операции вставки, удаления, замены и перестановки символа соответственно.

В результате работы алгоритма Дамерау-Левенштейна составляется матрица преобразования символов, где ее значения - это расстояния для преобразования символов. Итоговым значением преобразования одной строки в другую будет элемент с максимальными индексами. Например, расстояние между словами «дагестан» и «арестант» будет равно 3. Данный алгоритм требует довольно значительных затрат памяти и времени выполнения. Однако он может быть упрощен путем сокращения количества вычислений. Сравнение строк может происходить с учетом морфем, так могут подвергаться трансформации только суффиксы слов или окончания. Другим вариантом снижения затрат может быть использование ограничений по количеству сравнений.

Ключевым недостатком данных методов является время выполнения поиска. При работе с мобильными устройствами время является важным критерием работы приложения. Однако данные методы можно оптимизировать за счет их комбинаций или внесения дополнительных условий, задаваемых пользователем или разработчиком.

ЛИТЕРАТУРА

1. Adobe Experience Design [Электронный ресурс]. – Режим доступа: <http://www.adobe.com/ru/products/experience-design.html>.
2. Google Material Design [Электронный ресурс]. – Режим доступа: <https://material.io/guidelines/#introduction-principles>.
3. Нечеткий поиск в тексте и словаре [Электронный ресурс]. – Режим доступа: <https://habrahabr.ru/post/114997>.

УДК 519.256

Студ. Д.А. Радиванович

Науч. рук. ст.преп. Ю.О. Герман

(кафедра информационных систем и технологий, БГТУ)

ПРИЛОЖЕНИЕ ДЛЯ ФИЛЬТРАЦИИ ЭЛЕКТРОННОЙ КОРРЕСПОНДЕНЦИИ

Спам — массовая рассылка сообщений различного содержания, пользователям, не выражающим желания их получать. Согласно определению «Лаборатории Касперского» спам — это анонимная массовая непрошенная рассылка. В этом определении важно каждое включенное в него слово. Анонимная: все страдают именно от автоматических рассылок со скрытым или фальсифицированным