



УДК 681.5

## МЕТОДЫ ПОВЫШЕНИЯ УСТОЙЧИВОСТИ СИСТЕМ С РАСПОЗНАВАНИЕМ РЕЧИ И ОЦЕНКА ВРЕМЕННЫХ ПАРАМЕТРОВ РЕЧЕВОГО ИНТЕРФЕЙСА

*В.А. Жожикашвили, А.Ю. Андрейчук,  
Н.В. Петухова, М.П. Фархадов*

*В статье анализируются вопросы устойчивой работы систем с использованием технологии распознавания речи. Приводятся методы повышения эффективности работы таких систем. Дается описание некоторых прикладных систем массового обслуживания, применяющих речевые технологии. Предложена методология оценки временных параметров речевого интерфейса.*

**Введение.** За два года, прошедших с момента опубликования статьи по вопросу применения технологии распознавания речи в автоматизированных системах массового обслуживания (АСМО) [1], в мире и в России произошел заметный прогресс в этой области. Он выразился как в успехах в области точности распознавания речи, так и в масштабах практического применения соответствующей высокой технологии. Новым достижениям и проблемам в области распознавания речи в АСМО и посвящена настоящая статья.

**Устойчивость систем с распознаванием речи.** Главным вопросом в области компьютерного распознавания речи является точность распознавания. Точность, превышающая 95%, практически не достигнута. Такой результат, по-видимому, потребует больших усилий и времени, если не возникнут новые идеи и принципы. Это побуждает разработчиков использовать обходные пути.

Введем в рассмотрение понятие "устойчивость систем с распознаванием речи". Это понятие шире, чем точность распознавания. Оно охватывает такие случаи, как сбой при передаче речи по каналам связи, обрыв связи по разным причинам, различные шумы и помехи. Устойчивость подразумевает способность системы продолжить работу, несмотря на противодействие внешней среды.

Остановимся на описании некоторых средств повышения устойчивости, которые мы назвали обходными путями. Правильней было бы назвать их

алгоритмическими или логическими средствами. Такие средства уже используются в практических разработках авторов.

Во многих приложениях в самом начале диалога бывает целесообразно запросить у клиента его контактный телефон и зафиксировать его. Такой простой прием оказывается весьма универсальным, ибо работает во всех перечисленных выше случаях. С помощью контактного телефона сотрудники обслуживающей системы продолжают и завершают диалог с клиентом. Эта способность системы не потерять клиента является очень важной, когда система функционирует в развитой рыночной среде и когда потеря клиента непосредственно приводит к потере дохода. Так обстоит дело в транспортных системах, например при заказе билетов на самолеты или заказе такси.

Еще один логический механизм повышения устойчивости используется в тех точках диалога, где неправильное распознавание может привести к неприятным последствиям. Здесь может быть несколько ситуаций. Одна из них случается, когда в систему заранее заложено знание возможных ошибок. Так, ни один месяц не может иметь более 31 дня; время дня не может превышать 24 ч и т.д. Нельзя заказать такси на вчера или на 15 ч, если в момент заказа уже 16 ч, и много других случаев. Такие ошибки выявляются и корректируются программно.

Другая ситуация встречается тогда, когда ошибка логически не проявляется, например, если искажился номер контактного телефона, оставаясь в пределах 7 знаков, или номер дома клиента — заказчика такси. В этих случаях эффективно работает переспрос со стороны сервера.

Переспрос может быть неэффективным, если в диалоге встречаются "плохие" слова — с высокой вероятностью искажения типа: слово  $i$  → слово  $j$ . В этом случае может потребоваться переспрос несколько раз подряд. Если бы система знала все плохие слова, то при искажении типа  $i \rightarrow j$ , которое отличается высокой вероятностью, она автоматически угадала бы слово  $i$ . Поиск истинного слова происходил бы быстрее, чем при переспросах. Для этого необходимо, чтобы система в процессе эксплуатации собирала статистику ошибок.

Существуют и другие вероятностные методы исправления ошибок распознавания. Например, по аналогии с так называемым кодовым расстоянием в теории кодирования может быть введено в рассмотрение фонетическое расстояние между словами. Чем больше расстояние между парой  $i \rightarrow j$ , тем меньше вероятность перехода  $i \rightarrow j$  или  $j \rightarrow i$  в процессе распознавания. Достоинство этого метода в том, что он позволяет развить математические подходы и численно оценивать вероятность искажения. Эффективность этого метода сейчас изучается.