

УДК 004.233.5, 004.5

Д.И. Самаль

## СПОСОБ ВВОДА ТЕКСТА В ВЫЧИСЛИТЕЛЬНЫЕ УСТРОЙСТВА С РЕДУЦИРОВАННОЙ КЛАВИАТУРОЙ

*Предлагается альтернативный способ ввода текста в вычислительные устройства с редуцированной клавиатурой, который не требует внесения никаких дополнительных аппаратных изменений в конструкцию устройств. Он может быть реализован как для мобильных телефонов, смартфонов, так и для банкоматов, телевизионных приемников и других устройств. Способ позволяет вводить символы языков с алфавитами до 40 букв и оптимизирован для каждого конкретного языка – учитывает частоты встречаемости символов. Приводится сравнение способа с широко распространенными Multitar и T9™.*

### Введение

Ввод текста в цифровые устройства с редуцированной клавиатурой (мобильные телефоны, наладонные компьютеры и т. п.) вызывает сложности у пользователей, так как назначение для каждого символа алфавита отдельной клавиши не всегда представляется возможным – физических клавиш, как правило, всегда меньше, чем символов требуемого языка. Например, стандартная цифровая клавиатура мобильного телефона имеет 12 физических клавиш, дополнительно могут быть задействованы три-пять клавиш (в зависимости от конкретной модели). В подобных случаях возникает неоднозначность сопоставления физических клавиш устройства и символов алфавита (как букв, так и символов пунктуации, например «'» и «^» для французского языка или «¿» и «¡» для испанского).

Актуальность проблемы ввода с помощью редуцированных клавиатур подтверждается регулярным появлением на рынке дополнительных физических устройств, специально разработанных для подключения к функционально ограниченным устройствам – мобильным телефонам, коммуникаторам и т. п. Как правило, они содержат полноалфавитную клавиатуру и позволяют достаточно эффективно вводить и редактировать текст. Недостатки этих устройств заключаются в высокой стоимости, дополнительном весе и общем размере. Широкого распространения они не получили.

В статье рассматриваются существующие способы ввода текста в устройства с редуцированной клавиатурой и описывается разработанный альтернативный способ, превосходящий по эффективности и удобству прототипы.

### 1. Состояние проблемы

Вопрос улучшения системы ввода символов с помощью редуцированной клавиатуры, т. е. имеющей всего 12 клавиш, был поднят до наступления эры мобильной телефонии. Еще в 80-х гг. прошлого века проводились исследования эффективного распределения символов по клавишам телефона для передачи сообщений по телефонным каналам связи, тогда еще аналоговым. Их результаты обобщены в статье Дж. Арнотта и М. Джаведа [1]. В описанных в статье экспериментах раскладки символов латинского алфавита анализировались с точки зрения как оптимизации числа нажатий клавиш, так и повышения помехоустойчивости при разборе сообщения.

Исследования выявили, что при наборе текстов раскладка символов по алфавиту нерациональна и для каждого языка оптимальной будет своя собственная раскладка, составленная с учетом частот встречаемости символов в словах языка [1], т. е. буквы на клавишах следует поменять местами так, чтобы наиболее часто встречающиеся вводились наименьшим числом нажатий, чем наиболее редкие. Однако следует учитывать, что одни и те же буквы, например латинские, в различных языках используются с разными частотами, поэтому для каждого из языков наносить символы на клавиши следует по-разному. Таким образом,

оптимальная раскладка символов для немецкого языка, например, будет отличаться от оптимальной раскладки для французского или итальянского, несмотря на определенную схожесть алфавитов.

Раскладка символов в соответствии с английским алфавитом, которая фактически является стандартом для всех современных телефонов, была признана одной из самых неэффективных для ввода текста [1]. Следует отметить, что изначально символы на стационарных телефонных аппаратах появились в 20-х гг. прошлого века в Америке. Они предназначались для облегчения запоминания шестизначных телефонных номеров. Часть номера или весь номер полностью кодировался с помощью букв, при этом по техническим причинам клавише 1 символы алфавита не присваивались (1 использовалась для вызова оператора), буква Q не использовалась из-за сходства с 0. Иными словами, изначально раскладка по английскому алфавиту была введена для решения совершенно иной задачи и для кодирования слов она не предназначалась. Помимо исторической значимости, единственный плюс, который в результате и перевесил все минусы, заключался в универсальности этой раскладки для всех языков, использующих латиницу. Другими словами, для английского, французского, немецкого и им подобных языков она одинакова неудобна, однако перерисовывать символы на клавишах телефонного аппарата, в зависимости от места продажи телефона, не требовалось, что в результате и оказалось весомее остальных неудобств.

В настоящее время для ввода текста в мобильные телефоны наиболее широко используются два способа: Multitar, основанный на последовательном переборе всех назначенных клавише символов, и предиктивный ввод, реализованный в виде системы T9™. Данная система основана на поиске по словарю слов, которые содержат символы, назначенные нажатым клавишам, и выборе наиболее вероятного слова в зависимости из контекста или иных признаков. В обоих способах используются раскладки символов в соответствии с алфавитными порядками языков.

При использовании способа Multitar (в переводе с английского языка дословно означает «многонажатие») для ввода одного символа требуется один или несколько раз нажать физическую клавишу, поставленную в соответствие вводимому символу. Количество нажатий соответствует позиции вводимой буквы в ряде символов, поставленных в соответствие данной клавише. Такой способ де-факто является стандартным и широко используемым для мобильных телефонов. Его основными недостатками являются большое количество нажатий на физические клавиши при вводе одного символа и, как следствие, низкая скорость ввода, а также неудобство в использовании.

Самой известной из технологий ввода, устраняющих неоднозначность ввода символов на уровне слов, является T9™. Устройство реализации T9™ и способ ввода текста с ее помощью защищены патентами [2–5]. Основой технологии T9™ является автоматическая проверка по словарю выбранного пользователем языка возможных комбинаций всех букв, назначенных нажатым пользователем клавишам, с целью определения наиболее вероятного вводимого слова, т. е. система пытается выбрать требуемое слово из нескольких возможных вариантов, которые соответствуют различным комбинациям символов, нанесенных на нажатые клавиши. С помощью T9™ удастся облегчить ввод текста, состоящего из хранящихся в словаре слов, и повысить скорость ввода по сравнению с Multitar.

Технологии T9™ присущи определенные недостатки. Так, например, в случае неоднозначного определения системой вводимого слова пользователю требуется выбрать один из нескольких найденных программой вариантов. Для коротких слов, состоящих из трех-пяти букв, количество вариантов может превышать количество букв в слове. Если же вводимого пользователем слова нет в списке предложенных вариантов и, соответственно, в словаре, то пользователю предлагается ввести слово с помощью стандартного способа – последовательного ввода букв слова путем частичного или полного перебора всех назначенных клавише символов, т. е. с помощью способа Multitar.

В процессе увеличения размера доступной памяти портативных цифровых устройств ограничения на физический размер словаря системы T9™ могут быть сняты частично или полностью, однако при этом будет возрастать время на перебор увеличивающегося в арифметической прогрессии количества вариантов слов, которые могут быть составлены из всего набора симво-

лов, поставленных в соответствие нажатым пользователем клавишам. Дополнительным недостатком является невозможность ввода большинства из имен собственных, аббревиатур, технических терминов и других категорий текстовой информации. Ввод статистически надежных паролей с помощью данного способа также невозможен. Ограничения объема памяти в настоящее время еще влияют на количество одновременно доступных для ввода языков, однако даже если в устройстве будет достаточно памяти для хранения словарей всех языков мира, воспользоваться более чем двумя из них будет невозможно, так как на 12 клавиш затруднительно нанести символы алфавитов более чем двух языков. Чтобы задействовать систему для ввода текста на третьем языке, пользователь должен будет прибегнуть к помощи сменной клавиатуры либо выучить наизусть распределение символов алфавита требуемого языка по клавишам устройства, что не всегда возможно.

## 2. Альтернативный способ ввода

Следует отметить, что описанные выше популярные способы ввода были разработаны и стали широко использоваться относительно недавно – 10-20 лет назад, однако за последние пять лет устройства, для которых они используются чаще всего, – мобильные телефоны (в дальнейшем – телефоны) – сильно изменились с технической точки зрения. Прежде всего, у современных телефонов появился большой цветной экран, в то время как еще пять лет назад 90 % телефонов имели монохромный экран, отображающий лишь четыре-пять строк текста. Значительно повысились вычислительная мощность и объем памяти устройств. Указанные изменения позволяют перенести символы, соответствующие клавишам, с самих клавиш на экран, т. е. принципиально изменить организацию ввода.

*Отображение символов на экране* является основополагающей идеей предлагаемого способа. Для ввода символов предлагаемым способом используются всего 8 из 12 клавиш стандартной телефонной клавиатуры. Остальные клавиши предназначены для иных целей: изменения регистра и режима ввода, вставки символов пунктуации и т. п. Кроме того, одна клавиша отведена для наиболее часто встречающегося символа большинства языков – «пробела». В случае моделирования виртуальной клавиатуры на экране телефона 8 клавиш занимают на экране существенно меньше места, чем 12. Таким образом, 8 клавишами вводятся от 26 (в случае английского языка) до 33 (в случае русского) букв алфавита. Соответственно на каждую клавишу приходится по три-четыре английских буквы или по четыре-пять русских. Если каждая клавиша на экране в один момент времени отображает только один символ, то какая-то отдельная клавиша должна быть задействована для смены (переключения) символов на изображениях клавиш.

В результате экспериментов было принято решение размещать клавиши на экране по его периметру. Помимо удобства восприятия (клавиши разнесены по экрану), данный вариант расположения изображений клавиш позволил в случае необходимости отображать более восьми изображений клавиш – места по периметру экрана достаточно, а область вывода текста не меняет своей формы и экранных координат. Для ввода символов были определены клавиши редуцированной клавиатуры, располагающиеся по кругу относительно клавиши «5». Для телефона это клавиши «1»–«4» и «6»–«9». Клавиша «5», находящаяся в центре круга, стала отвечать за смену групп символов на изображениях клавиш на экране (рис. 1).

*Частоты встречаемости символов.* Как упоминалось выше, в статье [1] было показано, что частоты встречаемости символов при организации ввода символов посредством редуцированной клавиатуры следует учитывать обязательно. Соответственно каждый язык будет иметь индивидуальную раскладку символов. Частоты встречаемости символов для основных европейских языков приведены в табл. 1. Вторым базовым принципом нового способа ввода является *учет частот встречаемости символов каждого конкретного языка.*

*Разбиение клавиш по группам.* Очевидно, что для обеспечения эффективности ввода символы, которые чаще остальных встречаются в текстах, должны вводиться одним нажатием. В результате экспериментов была принята следующая схема разделения символов на группы. Все буквы алфавита и некоторые символы пунктуации разделены на три группы. К символам первой группы (наиболее часто присутствующим на экране) отнесены восемь символов с макси-

мальной частотой встречаемости в реальных текстах, второй группы – восемь букв со средними частотами, а третьей – все оставшиеся символы алфавита и наиболее востребованные символы пунктуации. При этом символы первой группы (или круга) всегда вводятся одним нажатием – после ввода любого символа из другой группы производится автоматическое переключение к символам первой группы.

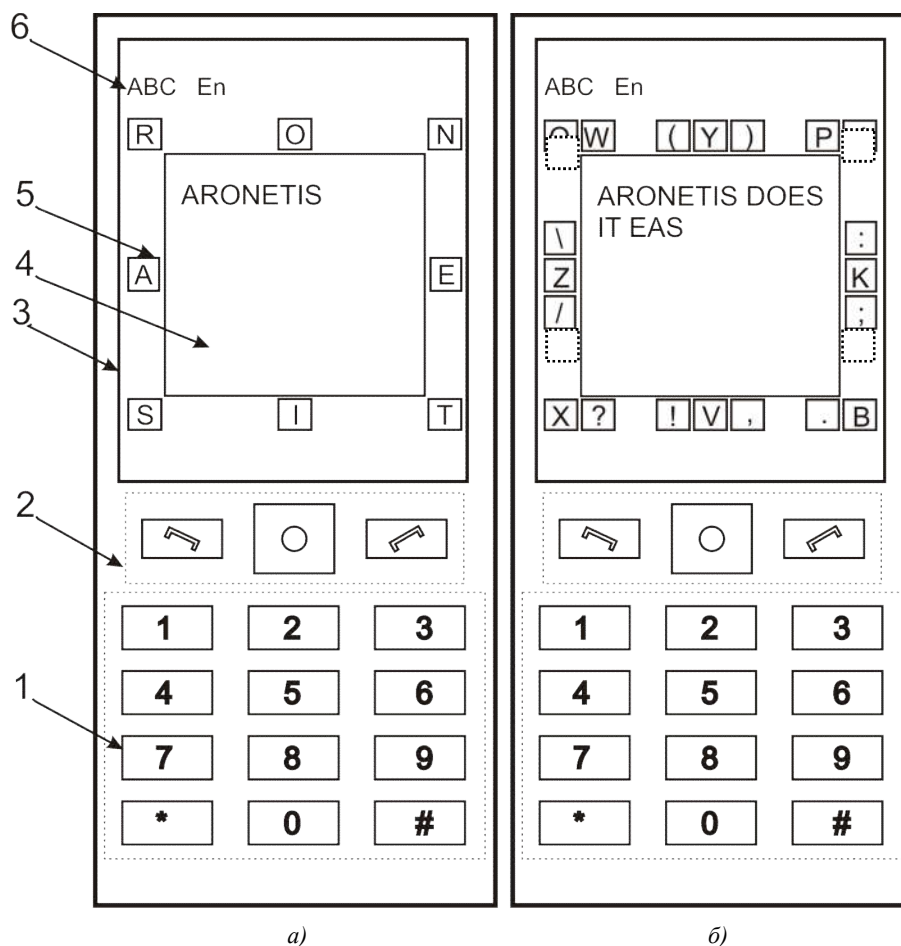


Рис. 1. Способ ввода текста в мобильные телефоны (на примере раскладки английского языка): а) первая группа символов; б) третья группа символов. Здесь 1 – редуцированная клавиатура устройства, 2 – область вспомогательных клавиш телефона, 3 – дисплей, 4 – область отображения набранного текста, 5 – символы активной группы, 6 – строка статуса (регистр, язык)

Ввод буквы из второго круга потребует двух нажатий: одного нажатия на клавишу «5» для переключения группы и нажатия на клавишу, соответствующую нужному символу.

Третья группа из-за своей многочисленности (она может включать в себя до 24 символов) требует четырех нажатий для ввода любого из своих символов – при нажатии дважды на «5» все символы данного круга высвечиваются на экране по его периметру, они визуально сгруппированы по тройкам (рис. 1, б). На данный момент в определенных группах одно из мест зарезервировано (на рис. 1, б показаны пунктиром), что позволяет использовать способ на устройствах с относительно небольшими по вертикали экранами. Каждая подгруппа соответствует своей клавише, и после нажатия на нее на экране остаются только символы данной подгруппы, вводимые одним нажатием клавиши.

Данный способ организации ввода редко встречающихся символов оказался наиболее удобным из всех рассматривавшихся. С его помощью можно осуществлять ввод на языках, не поддерживаемых производителями мобильных телефонов (например, белорусском), чей алфавит составляет до 40 символов. Если добавить еще две клавиши (например, «\*» и «#»), то без коренной ломки данного принципа количество символов в раскладке можно расширить до 50.

Таблица 1

## Частоты встречаемости букв алфавита в текстах языков

Русский		Английский		Французский		Немецкий		Испанский		Итальянский		Польский	
Буква алфавита	Частоты встречаемости	Буква алфавита	Частоты встречаемости	Буква алфавита	Частоты встречаемости	Буква алфавита	Частоты встречаемости	Буква алфавита	Частоты встречаемости	Буква алфавита	Частоты встречаемости	Буква алфавита	Частоты встречаемости
Пробел	0,2005	-/-	Нет данных	-/-	Нет данных	-/-	Нет данных	-/-	Нет данных	-/-	Нет данных	Пробел	0,1419
О	0,0764	Е	0,1286	Е	0,1776	Е	0,1918	Е	0,1415	І	0,1204	А	0,0756
Е	0,0732	Т	0,0972	S	0,0823	N	0,102	А	0,129	Е	0,1163	Е	0,068
А	0,0629	А	0,0796	А	0,0768	І	0,0821	О	0,0884	А	0,1112	О	0,066
И	0,0577	І	0,0777	N	0,0761	S	0,0707	S	0,0764	О	0,0892	І	0,0647
Т	0,0549	N	0,0751	Т	0,073	R	0,0701	І	0,0701	N	0,0768	N	0,0517
Н	0,049	R	0,0683	І	0,0723	Т	0,0586	R	0,0695	Т	0,0707	R	0,0435
Р	0,0459	О	0,0662	R	0,0681	А	0,0552	N	0,062	R	0,0656	Z	0,0399
С	0,0404	S	0,0662	U	0,0605	H	0,0502	L	0,0552	L	0,0595	W	0,0397
В	0,0355	H	0,0539	L	0,0589	D	0,0491	D	0,0467	S	0,0481	T	0,0395
П	0,033	D	0,0401	О	0,0534	U	0,0422	C	0,0442	C	0,0411	S	0,0368
К	0,0302	L	0,0351	D	0,036	G	0,036	T	0,0436	D	0,0354	C	0,0336
Л	0,0299	C	0,0284	C	0,0332	L	0,0348	U	0,04	U	0,0309	Y	0,0333
М	0,0275	F	0,0262	P	0,0324	C	0,0294	P	0,0326	P	0,0266	P	0,03
Д	0,0265	U	0,0248	M	0,0272	О	0,0214	M	0,0255	M	0,0265	K	0,029
У	0,0222	M	0,0243	Q	0,0134	F	0,0196	Q	0,0155	G	0,0173	D	0,0277
Я	0,0153	G	0,0199	V	0,0127	M	0,0169	Y	0,0105	V	0,0167	M	0,0263
Ы	0,0143	P	0,0181	G	0,011	B	0,0156	B	0,0103	Z	0,0124	U	0,0231
Ь	0,0138	W	0,018	F	0,0106	W	0,0138	G	0,01	F	0,0115	L	0,0216
З	0,0133	B	0,016	B	0,008	K	0,0133	H	0,0091	B	0,0107	J	0,0189
Й	0,0125	Y	0,0152	H	0,0064	Z	0,0117	F	0,007	H	0,0083	G	0,0121
Б	0,0114	V	0,0115	X	0,0054	V	0,0084	V	0,0067	Q	0,0048	B	0,0112
Ч	0,0094	K	0,0041	Y	0,0021	P	0,0054	Z	0,0031	J	0	H	0,0089
Г	0,0083	Q	0,0017	J	0,0019	J	0,0016	J	0,0024	K	0	Ą	0,0083
Ю	0,0081	X	0,0017	Z	0,0007	Q	0,0001	X	0,0007	W	0	Ł	0,008
Ж	0,0079	J	0,0016	K	0	X	0	K	0	X	0	Ę	0,0077
Х	0,0048	Z	0,0005	W	0	Y	0	W	0	Y	0	Ż	0,0069
Щ	0,0042											Ó	0,0067
Ф	0,0036											F	0,0052
Ш	0,0026											Ś	0,0052
Э	0,0023											Ć	0,0048
Ц	0,0021											X	0,0014
Ъ	0,0003											V	0,0012
												Ń	0,001
												Ż	0,0005
												Q	0,0002

Примечание: для русского [6] и польского [7] языков указаны вероятности с учетом пробельного символа, для остальных языков вероятности указаны без учета пробелов и символов, не входящих в латинский алфавит [8].

Распределение букв в каждой группе. Третьей идеей, на которой базируется способ, является размещение символа алфавита на экране мобильного устройства в соответствии с его положением на клавиатуре стандартного персонального компьютера. Это означает, что если, например, символ «Ч» на клавиатуре персонального компьютера располагается в левом нижнем углу относительно центра клавиатуры, то и на экране мобильного телефона его местоположение должно быть в левом нижнем углу. Примеры матриц раскладок для различных языков показаны на рис. 2.

Белорусский	Украинский	Итальянский	Испанский
Й Ц Г Ш Ё Э Х У О Л Ё Р К Н П В А Е Д Ж Ф С И Ы Э Я М Т Ч ? Ь Б , . Ю	Й Ц Г Ш Щ Э Х У К Л Ё Р О Н І Ф Е А І Д Ж ? В И Т Ё С М П Я Ч , Ь . Ю Б	Q W Y F È É À S U P J R O N Z D A E G Ù Ç L I T K C V M X ? ! Н , . В	F W Ñ G Ç Н J Q T U ¿ R O N Z D A E P K i S I L C Y M X ? ! V , . В
Русский	Немецкий	Польский	Французский
Й Ц Г Ш Щ Э Х У К Л Ё Р О Н Ъ Ы В А Е Д Ж Ф С И Т Э Я М П Ч ? Ь Б , . Ю	W Ö Ä Z Ü В Р F O U ! R H N : Q D A E L K Y S I T J C G M X ? ; V , . В	A G Ż U Z Ó J W D P F R O N Ł H S A E K L Ś Z I T Ę Y C M Ć X V B , . Ń	F Y é G è ç à Q O P E R T ^ Z D A U L ù W S N I " C V M X H J V B K , .
Шведский	Английский	Раскладка «123»	Раскладка SMB
Q W Y U J П F H O ^ R L N Å Z D A E K Ä " S I T Ö V G M X ? ! C , . В	Q W ( Y ) J P F H U ^ R O N : Z D A E L K / S I T ; C G M X ? ! V , . В	· Σ ( Δ ) { } + - = ≠ 7 8 9 [ = ± 4 6 * ⊗ ≈ 1 2 3 ] / , . + ⊕ × ∇ ≤ ≅ ≥	№ _ ^ & * { } # \$ % ~ ! ( ) [ ] ¥ @ ` : € ; - " / - ] < > _ + =   ° < > \

Рис. 2. Примеры раскладок. Жирным шрифтом выделены символы первой группы (в центре матрицы). Символы второй группы расположены по периметру первой. Третья группа занимает крайние клетки матрицы. Пунктирной линией выделены места (по четыре в каждой из раскладок), зарезервированные для дополнительных символов

### 3. Сравнение способов ввода

Ниже приведены результаты экспериментов ввода текстов с помощью разработанного способа. В табл. 2 даются максимальные теоретические значения эффективности предлагаемого способа и способа Multitar, т. е. предполагается, что пользователь не будет ошибаться при наборе символов. Для вычисления максимальной эффективности разработанного способа использовалась формула

$$F = 1*(Sum(G1)) + 2*(Sum(G2)) + 4*(Sum(G3)),$$

где Sum(G1) – сумма частот встречаемости символов первого круга раскладки языка, Sum(G2) – второго круга, Sum(G3) – третьего круга, коэффициенты 1, 2 и 4 соответствуют количеству нажатий клавиш для ввода одного символа соответствующей группы.

Таблица 2

Максимальная эффективность ввода (нажатий/символов)

Язык ввода	Предлагаемый способ	Multitar	Язык ввода	Предлагаемый способ	Multitar
Русский	1,5768	2,0867103	Испанский	1,4065	2,2208
Английский	1,5179	2,1632	Итальянский	1,3857	2,2373
Французский	1,4055	2,219	Польский	1,7146	2,04721
Немецкий	1,4591	2,194	Шведский	1,4509	1,8187

*Примечание.* 1. Максимальные эффективности вычислены по данным табл. 1: для русского, польского и шведского языков с учетом пробелов и всех символов алфавита, для остальных языков без учета пробелов и символов, не содержащихся в латинском алфавите. Символы пунктуации не учитывались. 2. При вычислении эффективности ввода с помощью Multitar для всех языков, за исключением русского, использовалась алфавитная раскладка английского языка (де-факто – стандартная для мобильных телефонов).

Вычисление максимальной эффективности способа Multitar производилось в соответствии со следующим выражением:

$$F = \sum_{i=1}^n k_i \times p_i,$$

где  $i$  – порядковая позиция символа в табл. 1, начиная с верха таблицы;  $n$  – общее количество символов в алфавите;  $k_i$  – количество нажатий, требуемых для ввода данного символа способом Multitar;  $p_i$  – частота встречаемости символа.

Для сравнения способов был выбран текст, традиционно используемый в соревнованиях на скорость набора SMS: «*Острозубые пираньи рода Сerrasальмус и Пигоцентрус являются наиболее опасными пресноводными рыбами в мире. Но в действительности они редко нападают на человека*». Результаты сравнения необходимого количества нажатий для ввода текста с помощью названных способов представлены в табл. 3. Английский вариант указанного текста следующий: «*The razor-toothed piranhas of the genera Serrasalmus and Pygocentrus are the most ferocious freshwater fish in the world. In reality they seldom attack a human*». Итоговые результаты сравнения способов при вводе английского варианта приведены в нижней строке табл. 3. Полная таблица результатов ввода английского текста дана в работе [9].

Анализ результатов подтверждает следующий факт: разработанный способ эффективнее известных на данный момент методов ввода текстов в мобильные телефоны. Multitar проигрывает предлагаемому способу по эффективности при любых вариантах вводимого текста.

Системы предиктивного ввода T9™ и iTap™ могут быть эффективнее рассматриваемого способа для простых текстов со средней длиной слова пять-шесть символов, однако в случае коротких (два-три символа) слов количество вариантов в словаре, соответствующих нажатым клавишам, может превышать количество букв в слове и в подобных случаях разработанный способ будет эффективнее. Аналогичная ситуация и со словами, состоящими из большого количества букв, – система T9™ обрабатывает их не всегда корректно.

Слова, не содержащиеся в словаре, при попытке их набора с помощью T9™ приводят к большим затратам усилий, чем при использовании даже Multitar, не говоря уже про предлагаемый способ. Вставка символов пунктуации – дополнительный выигрыш предлагаемого способа.

Таблица 3

Количество нажатий при вводе заданной фразы

Вводимое слово	Количество нажатий при вводе с помощью T9 (SE k500i)	Количество нажатий при вводе с помощью предлагаемого способа	Количество нажатий при вводе с помощью Multitap
Острозубые	9 (слово не найдено в словаре) + 6 (удаление лишнего и выбор «Остро») + 5 («зубые») = 20	1 («O») + 1 (shift) + 19 = 21	29
(пробел)	1	1	1
пираньи	7 («мирамял») + 4 (удаление лишнего и выбор «пира») + 7 («ньи») = 18	11	11
(пробел)	1	1	1
рода	4	5	6
(пробел)	1	1	1
Серрасальмус	1 (shift) + 8 («Террасал» и слово не найдено в словаре) + 9 (удаление лишнего и выбор «Сер») + 5 («ра») + 6 («са») + 4 («ль») + 16 («мус») = 49	1 (shift) + 1 («C») + 1 (shift) + 17 = 20	1 (shift) + 22 = 23
(пробел)	1	1	1
и	1	1	1
(пробел)	1	1	1
Пигоцентрус	1 (shift) + 6 («Оканч» и слово не найдено в словаре) + 16 (удаление лишнего и выбор «Пиг») + 2 («о») + 5 («цен») + 7 («трус») = 37	1 (shift) + 2 («I») + 1 (shift) + 17 = 21	1 (shift) + 29 = 30
(пробел)	1	1	1
являются	8	16	27
(пробел)	1	1	1
наиболее	8	12	18
(пробел)	1	1	1
опасными	8	13	18
(пробел)	1	1	1
пресноводными	8 («пресповн») + 4 (удаление лишнего и выбор «пресно») + 7 («годными») + 3 (удаление и выбор «водным») + 2 («и») = 24	22	29
(пробел)	1	1	1
рыбами	6	13	10
(пробел)	1	1	1
в	1	2	3
(пробел)	1	1	1
мире	4	5	5
. (точка)	5	4	5
(пробел)	1	1	1
Но	2	1 (shift) + 1 («H») + 1 (shift) + 1 = 4	5
(пробел)	1	1	1
в	1	2	3
(пробел)	1	1	1
действительности	16	25	35
(пробел)	1	1	1
они	3	3	6
(пробел)	1	1	1
редко	7	7	10
(пробел)	1	1	1
нападают	8	13	16
(пробел)	1	1	1
на	2	2	3
(пробел)	1	1	1
человека	8	14	23
. (точка)	4	4	4
<b>Всего нажатий</b>	<b>264</b>	<b>260</b>	<b>340</b>
<b>Итоговое количество нажатий при вводе английского варианта текста</b>	<b>260</b>	<b>237</b>	<b>334</b>

*Примечание.* В ходе эксперимента ввод с помощью T9™ слов, не содержащихся в словаре, потребовал больше нажатий, чем указано в таблице. Здесь приведены минимальные значения, которых удалось добиться после нескольких попыток.



### Заключение

По сравнению с Multitar, предлагаемый способ обеспечивает существенный математически доказуемый выигрыш в уменьшении среднего числа нажатий клавиш для ввода одного символа текста. По сравнению с системами T9™, прямой выигрыш в количестве нажатий не столь значителен, но при этом достигаются следующие преимущества над T9™:

- возможность набора слов, не содержащихся в словарях (имена, фамилии, адреса, аббревиатуры, термины, пароли и т. п.);
- возможность набора сообщения, содержащего слова из разных языков, причем произвольного количества языков одновременно;
- возможность ввода символов различного назначения (например, математических или специальных символов транскрипции);
- легкость переключения и подключения дополнительных языков;
- независимость способа от символов, нанесенных на клавиши телефона;
- нечувствительность системы к орфографическим ошибкам пользователя;
- набор кириллического текста латинскими символами либо автоматическая конвертация набранного кириллицей текста в транслит;
- представление всего спектра unicode-символов;
- ввод сообщений на любом языке, имеющем алфавит до 40 (или до 50 – в расширенном варианте) символов;
- создание индивидуальных раскладок для ввода устойчивых к взлому паролей, причем могут использоваться все unicode-символы;
- возможность ввода текста на практически любом программируемом устройстве с дисплеем и редуцированной клавиатурой (содержащей от шести и более клавиш), например в банкоматах, телевизорах (в комплекте с пультом дистанционного управления) и т. п.

С теоретической точки зрения система T9™ устраняет неоднозначность ввода редуцированной клавиатуры на уровне слов, а предлагаемый способ – неоднозначность на уровне символов. Следует отметить, что в статье [1] делается вывод о перспективности устранения неоднозначности ввода именно на уровне символов, по сравнению с устранением неоднозначности на уровне слов. Данный вывод подтверждается принципиальной возможностью набора текста разработанным способом «вслепую», т. е. не глядя ни на экран, ни на клавиатуру.

По оценкам бета-тестеров, разработанный способ удобнее и функциональнее названных выше систем. Новизна способа подтверждена тремя патентами [10–12]. С его помощью впервые с мобильного телефона было послано SMS-сообщение на белорусском языке.

### Список литературы

1. Arnott, J.L. Probabilistic character disambiguation for reduced keyboards using small text samples / J.L. Arnott, M.Y. Javed // AAC Augmentative and Alternative Communication. – 1992. – 8 (1). – P. 215–223.
2. Reduced keyboard disambiguating system: пат. 6,307,549 США, МПК G06F 1300 / M.T. King, D.L. Grover, C.A. Kushler, Ch. A. Grunbock; заявитель Tegic Communications Inc.; опубл. 23.10.2001 // US Patent and Trademark Office. Official Gazette. – 2001. – № 43.
3. Reduced keyboard disambiguating system: пат. 6,307,548 США, МПК G06F 1300 / E.P. Flinchem, D.L. Grover, Ch.A. Grunbock, M.T. King, C.A. Kushler; заявитель Tegic Communications Inc.; опубл. 23.10.2001 // US Patent and Trademark Office. Official Gazette. – 2001. – № 43.
4. Reduced keyboard disambiguating system: пат. 6,011,554 США, МПК G06F 1300 / M.T. King, D.L. Grover, C.A. Kushler, Ch. A. Grunbock; заявитель Tegic Communications Inc.; опубл. 04.01.2000 // US Patent and Trademark Office. Official Gazette. – 2000. – № 1.
5. Система устранения неоднозначности редуцированной клавиатуры: пат. 2221268 Российская Федерация, МПК G06F3/023 / М.Т. Кинг, Д.Л. Грувер, К.А. Кушлер, Ч.А. Грунбок; заявитель Теджик Коммьюникейшнз Инк. (US); опубл. 10.01.2000 // Официальный бюл. «Изобретения. Полезные модели» / Федеральный институт промышленной собственности (ФИПС). – 2000. – № 1.

6. Анализ текстов // Statsoft [Электронный ресурс]. – М., 2007. – Режим доступа: <http://www.statsoft.ru/home/portal/exchange/textanalysis.htm>. – Дата доступа: 19.11.2007.
7. Letter frequency distribution for Polish // Ceti [Electronic resource]. – Варшава, 2007. – Mode of access: <http://www.ceti.pl/poleng/zasoby/dane/tstat/english.htm>. – Date of access: 19.11.2007.
8. Применение частотных характеристик текстов // Lg-web [Электронный ресурс]. – М., 2007. – Режим доступа: <http://lg--web.chat.ru/texts.html>. – Дата доступа: 19.11.2007.
9. Сравнение способов // Аронетис [Электронный ресурс]. – Минск, 2007. – Режим доступа: [http://www.aronetis.com/competition\\_en.html](http://www.aronetis.com/competition_en.html). – Дата доступа: 19.11.2007.
10. Способ ввода текста в электронно-вычислительные устройства: пат. 9377 Респ. Беларусь, МПК7 G 06 F 3/02 / Д.И. Самаль; заявитель Д.И. Самаль. – № a20041008; заявл. 29.10.2004; опубл. 30.06.07 // Афіцыйны бюл. / Нац. цэнтр інтэлектуал. уласнасці. – 2007. – № 2.
11. Матрица для ввода символов в электронно-вычислительные устройства: пат. 2958 Респ. Беларусь, МПК7 G 06 F 3/02 / Д.И. Самаль; заявитель Д.И. Самаль. – № u20050608; заявл. 14.10.2005; опубл. 30.08.06 // Афіцыйны бюл. / Нац. цэнтр інтэлектуал. уласнасці. – 2006. – № 3.
12. Способ ввода текста в электронно-вычислительные устройства: пат. 2 304 301 Российская Федерация, МПК7 G 06 F 3/02 / Д.И. Самаль; заявитель Д.И. Самаль. – № 2005126907/09; заявл. 28.08.2005; опубл. 10.03.07 // Официальный бюл. «Изобретения и полезные модели» / Федеральный институт промышленной собственности (ФИПС). – 2007. – № 3.

Поступила 04.10.07

*Объединенный институт проблем  
информатики НАН Беларуси,  
Минск, Сурганова, 6  
e-mail: samal@newman.bas-net.by*

**D.I. Samal**

### **METHOD FOR TEXT INPUT INTO COMPUTERS WITH REDUCED KEYBOARD**

A new method of text input to digital devices by reduced keyboard is presented. No any hardware changes or supplements of devices are needed. The method may be realized for mobile phones, smart-phones, ATM or cash machines, TV sets and other devices. It is suitable for any language with alphabets of up to 50 letters. It allows writing any proper names, technical terms, multilanguage messages in optimal way for each language. Especial advantage of the method is a possibility of typing statistically strong passwords. The method is much more effective than a standard way of typing such as «Multitap» and «T9™».