

A Framework for Secure Communication using Hindi for Web-based and Mobile Applications

Authors :

Dr. Saibal K. Pal

*Scientific Analysis Group,
DRDO, Delhi India*

Sarvesh Kumar

Sarvejeet Kumar

Mukesh Kumar

*Department of computer Science
University of Delhi, Delhi India*

Need for Translation and Security of Regional Language

- **Translation**

As we know India is a multilingual country, so communication takes place in many regional languages .

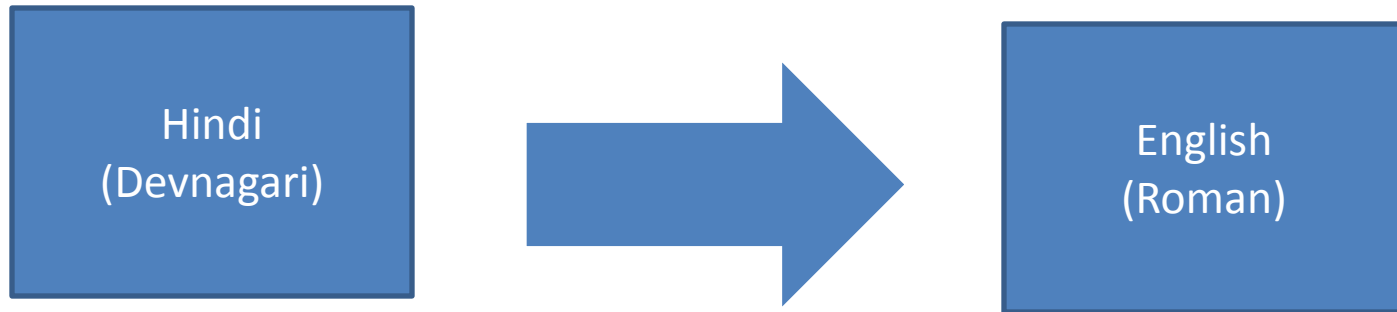
The translation of a regional language into corresponding English script will provide a way to connect people through their native languages.

- **Security**

Many a times it is required to secure a communication in regional language over the networks.

For example to secure an e-mail in a regional language.

Script Translation



How ??

Rule-based Script Translation

- DEVNAGARI
 - Phonetics based script
 - Words are largely written according to phonetics(pronunciation)
 - Possibility that a written word is pronounced differently is very less
 - Does not require memorizing spellings

These characteristics of Devnagari Script makes it EASY to convert it into Roman

Rule-based Script Translation (contd.)

- The pronunciation of Hindi(Devnagari) words is not continuous, it takes momentary stops in between.

e. g. सरगम (SARGAM)

This Characteristic of Devnagari Script makes it DIFFICULT to convert Devnagari into Roman.

Rules used in Translation

- **Anusvaar rule**

The occurrence of Anusvaar(ं) in a Hindi word produces two sounds 'ँ' and 'ं' depending upon the next consonant in the word .

If the next consonant is from 'प वर्ग' then Anusvaar produces sound of 'ं' otherwise sound of 'ँ'.

e.g. संबंध = SAMBANDH

- **Laghu swar and Deergh swar rule**

e.g.

कमला = *Kamla* OR *Kamala* (inappropriate but correct according to std. Hindi)

In this example there is a DEERGH SWAR on 'ल' hence 'a' is not placed between 'm' and 'l'.

Rules used in Translation (contd.)

- **Rearrangement of *vowel marks* (MATRA)**

Rearranging REF (')

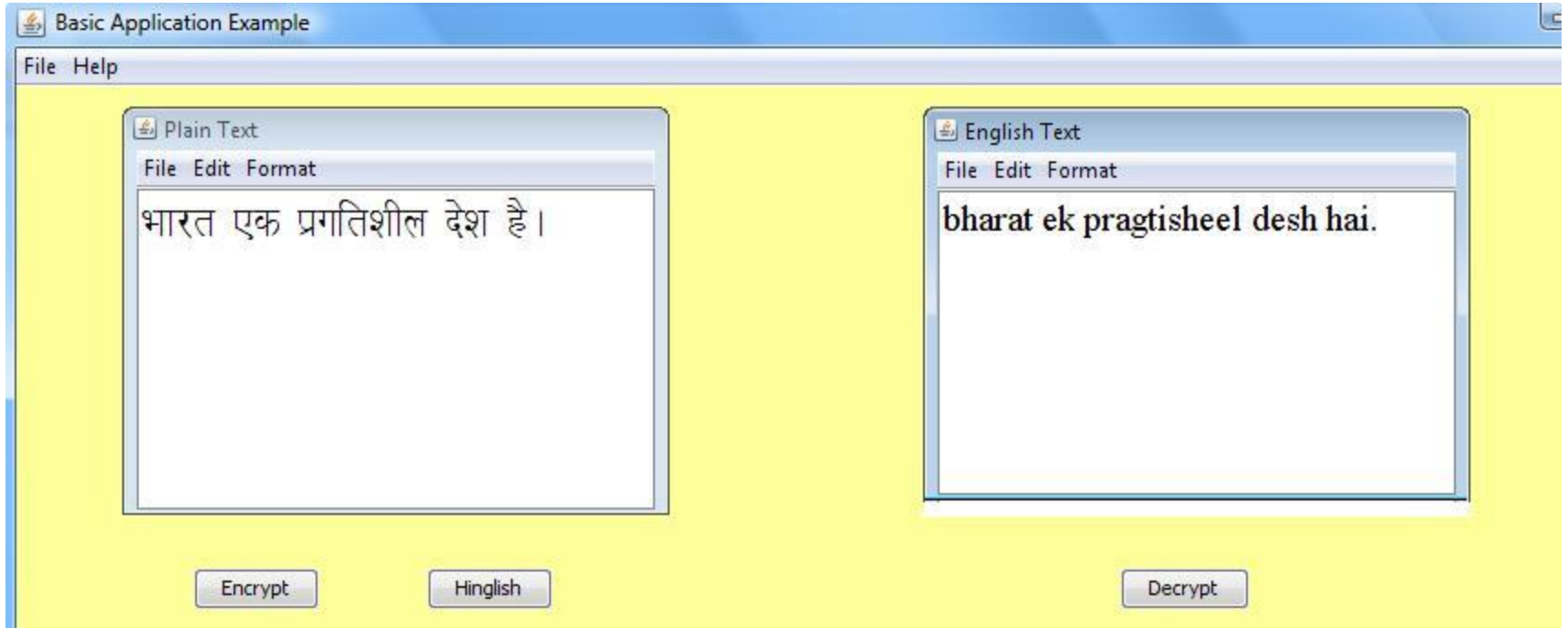
e.g. सर्वजन is written as स व ' ज न
 सर्वजन is pronounced as स ' व ज न (SARVJAN)

Rearranging (ि)

e.g. किताब is written as ि क त ा ब
 किताब is pronounced as क ि त ा ब (KITAB)

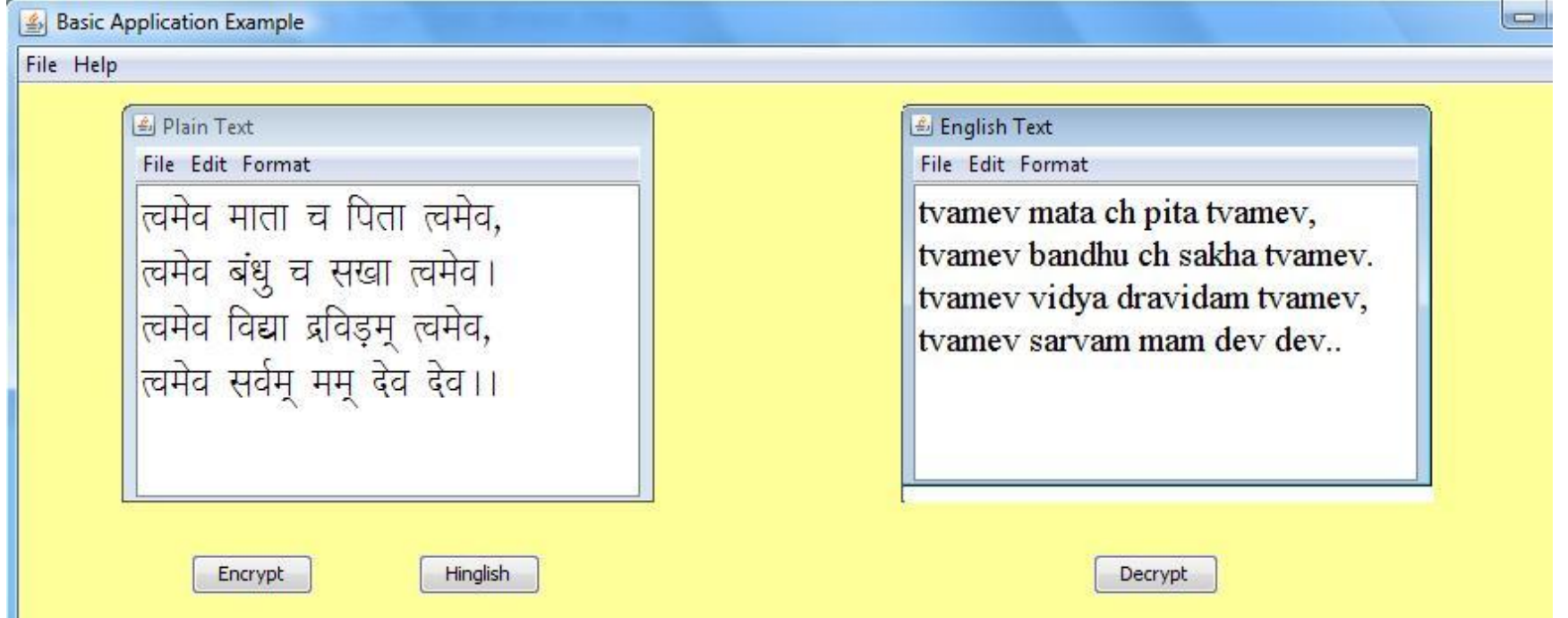
Here the Matra(ि) is placed at the position where it's pronunciation occurs.

Screen-shot



Hindi (Devnagari) to English (Roman)

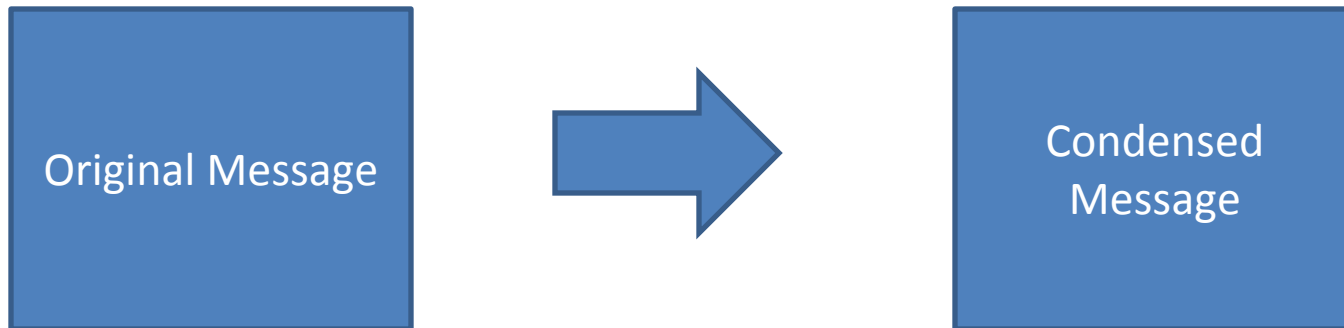
Screen-shot



Sanskrit (Devnagari) to English (Roman)

Rule-based Condensation of Message

- We can condense the message according to rules of texting languages as we do in ***SMSs***.



Rule-based Condensation of Message (contd.)

- Abbreviations and Symbols can be used for certain frequent words in the message.

e.g. **mob.** for **mobile**
@ for **at**
& for **and**
2day for **today**

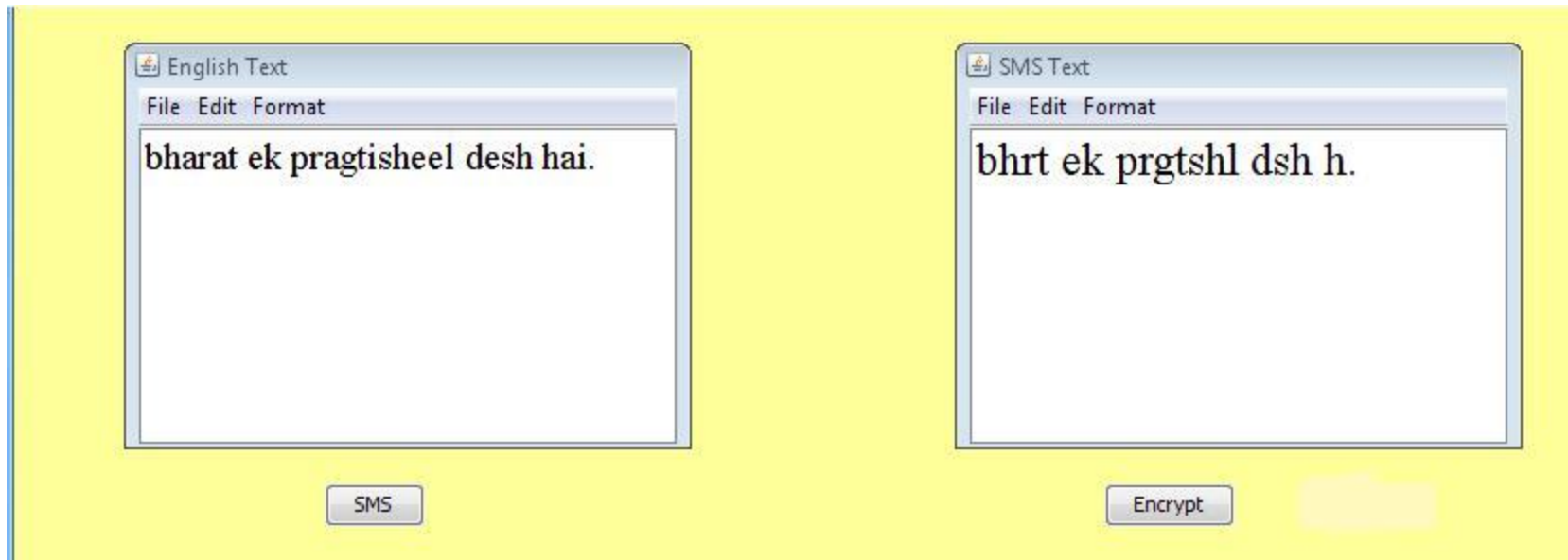
etc.

- Vowels can be omitted in a word in such a way that reduced word convey same meaning as that of original. Repetition of consonants can also be omitted.

e.g.

shortest → **shrtst**
good → **gd**
something → **smthng**
message → **mssg** → **msg**

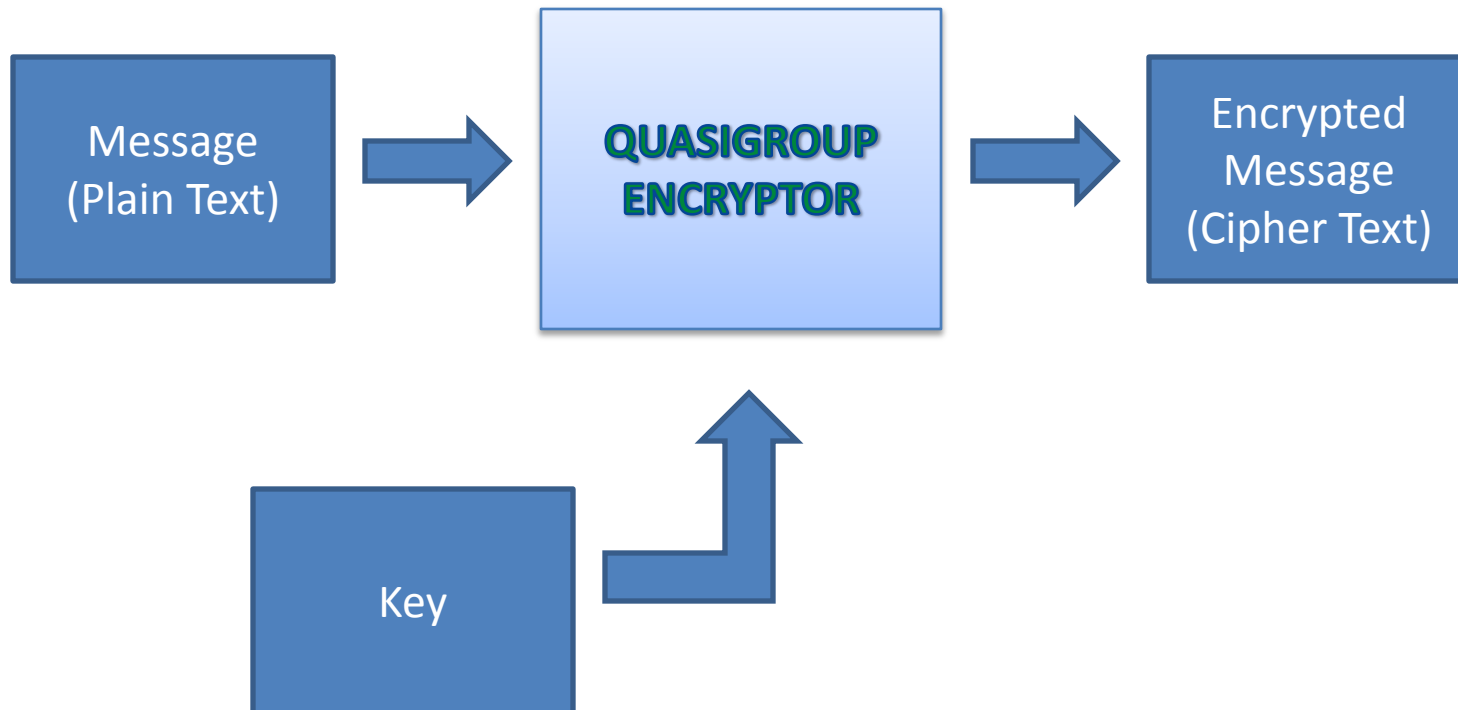
Screen-shot



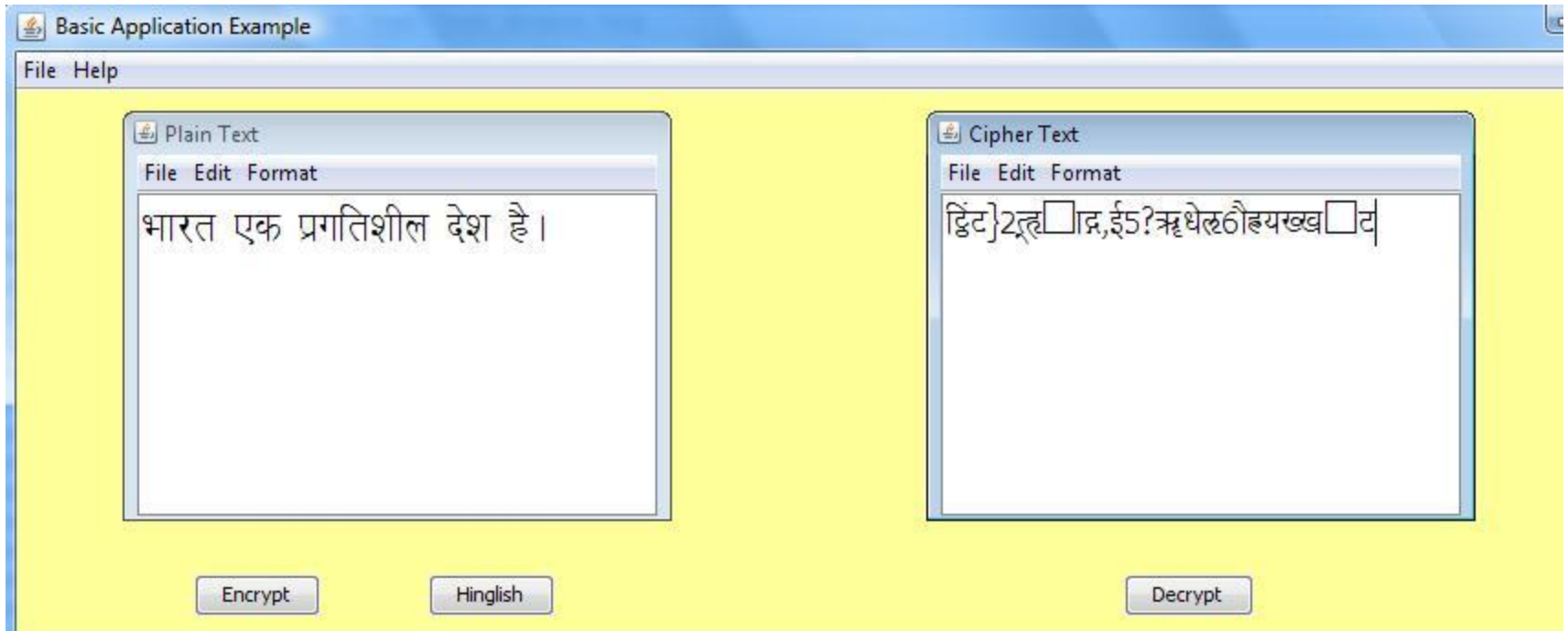
English to SMS

Securing communication over the WEB

- To ensure security of a message (*Hindi* or *English*), message is encrypted using **QUASIGROUP** encryption scheme. **QUASIGROUP** of order 256 is used.



Screen-shot



Encryption

Thank You

Contact:

Mobile : +919650944960

Email : stonesarvesh@gamil.com