Communication over the Internet

E-mail
MIME
Newsgroup
Netiquette
Instant Messenger
Internet Telephony

What is e-mail?

• A method of composing, sending, and receiving messages over electronic communication systems

• Started approximately in 1965, pre-dates to the Internet
  – on time-sharing mainframe computers
  – on networked computers

Internet e-mail addresses

• E-mail address format
  userID@hostname

• E-mail address is not unique for users
  One person can have many user IDs and host names.

• E-mail address need not reflect user identity.

How does Internet e-mail work?

• The From: e-mail address of an e-mail
  – may be faked
  – but one still should try not to allow other people to use one’s e-mail address

• The ARPANET computer network became a major cornerstone in the evolution of e-mail

• Multiple e-mail protocols existed
  – UUCP: Unix to Unix Copy Protocol (need to specify the route)
  – VNET: a corporate e-mail system from IBM

• Modern Internet e-mail
  – SMTP: Simple Mail Transfer Protocol
    From client to a "remote" host
    ASCII- and Line-based
  – POP3, IMAP:
    From host server to client

• The From: e-mail address of an e-mail
  – may be faked
  – but one still should try not to allow other people to use one’s e-mail address
Simple Mail Transfer Protocol

• Simple Mail Transfer Protocol (SMTP)
  – A simple, line- and ASCII-based protocol
    • One or more recipients
    • In most cases verified to exist
    • Then the message text is transferred (sent and received).
  – The defacto standard
    For e-mail transmission across the Internet

– Utilizes some Mail Transfer Agent (MTA)
  Also called Mail Server, Mail Exchange Server, or a SMTP server.
  e.g. sendmail (There are over 50 MTA implementations.)
  – Works best when both the sending and receiving machines are connected to the network all the time.

– The basic format of Internet e-mail is defined in RFC 2822
  
  http://www.faqs.org/rfcs/rfc2822.html
  
  • formats for text e-mail headers and body
  • rules pertaining to commonly used header fields
    – "To:"
    – "Subject:"
    – "From:"
    – "Date:"

– Cannot authenticate the sender and cannot take care of spamming.
  Too costly to revise.
  (New proposals are in the works.)

– Most SMTP servers today support
  Multipurpose Internet Mail Extension (MIME)
  for transmitting different types of data

Multipurpose Internet Mail Extensions

• Multipurpose Internet Mail Extensions (MIME)
  – Mechanisms for sending information other than line-based English language text in e-mail
    • Text in languages other than English using character encodings other than ASCII (7-bit ASCII)
    • 8-bit binary content such as files containing
      – images
      – sounds
      – movies
      – computer object programs
      – etc.
A standard for multi-part, multimedia on the Internet

- e-mail messages (SMTP)
- WWW hypertext documents (HTTP)

Mapping messages into and out of MIME format

Typically done automatically by an email client or by proprietary plug-ins.

Common MIME Types

Some examples:
- .html .htm - Hypertext markup language
- .txt - unformatted text
- .wav - Microsoft sound
- .au - Sun Microsystems sound
- .jpeg .jpg - Joint Photographic Experts Group
- .gif - Graphic Interchange Format
- .avi - Microsoft Audio Video Interleaved
- .mpeg .mpg - Movie Picture Experts Group

See Table 1.3 (p.42) in the textbook for more.
New types may be registered.

ASCII code for text

- Different computer file types are coded differently. However, they all end up as bits of 0's and 1's.
- Modern computers deal with 8 bits, a byte, as the basic group/unit to represent data.
- Coding text is the easiest, since there are only a limited number of different characters in a language.

A partial list of the ASCII code

<table>
<thead>
<tr>
<th>Dec</th>
<th>Oct</th>
<th>Hex</th>
<th>Binary</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>032</td>
<td>040</td>
<td>20</td>
<td>00100000</td>
<td>SP (Space)</td>
</tr>
<tr>
<td>033</td>
<td>041</td>
<td>21</td>
<td>00100001</td>
<td>! (exclamation mark)</td>
</tr>
<tr>
<td>034</td>
<td>042</td>
<td>22</td>
<td>00100010</td>
<td>&quot; (double quote)</td>
</tr>
<tr>
<td>035</td>
<td>043</td>
<td>23</td>
<td>00100011</td>
<td># (number sign)</td>
</tr>
<tr>
<td>036</td>
<td>044</td>
<td>24</td>
<td>00100100</td>
<td>$ (dolar sign)</td>
</tr>
<tr>
<td>037</td>
<td>045</td>
<td>25</td>
<td>00100101</td>
<td>% (percent)</td>
</tr>
<tr>
<td>038</td>
<td>046</td>
<td>26</td>
<td>00100110</td>
<td>&amp; (amperasnd)</td>
</tr>
<tr>
<td>039</td>
<td>047</td>
<td>27</td>
<td>00100111</td>
<td>’ (single quote)</td>
</tr>
<tr>
<td>040</td>
<td>050</td>
<td>28</td>
<td>00101000</td>
<td>( (left/opening parenthesis)</td>
</tr>
<tr>
<td>041</td>
<td>051</td>
<td>29</td>
<td>00101001</td>
<td>) (right/closing parenthesis)</td>
</tr>
<tr>
<td>042</td>
<td>052</td>
<td>2A</td>
<td>00101010</td>
<td>* (asterisk)</td>
</tr>
<tr>
<td>043</td>
<td>053</td>
<td>2B</td>
<td>00101011</td>
<td>+ (plus)</td>
</tr>
<tr>
<td>044</td>
<td>054</td>
<td>2C</td>
<td>00101100</td>
<td>, (comma)</td>
</tr>
<tr>
<td>045</td>
<td>055</td>
<td>2D</td>
<td>00101101</td>
<td>- (minus or dash)</td>
</tr>
<tr>
<td>046</td>
<td>056</td>
<td>2E</td>
<td>00101110</td>
<td>. (dot)</td>
</tr>
<tr>
<td>047</td>
<td>057</td>
<td>2F</td>
<td>00101111</td>
<td>/ (forward slash)</td>
</tr>
</tbody>
</table>

Conversion from binary

- Review decimal numbers, base 10:
  \[ 2893 = 2\times10^3 + 8\times10^2 + 9\times10^1 + 3\times10^0 \]

- Binary number, base 2, binary digits: 0 1
  e.g.: 01011111
  Binary bits, 0 or 1, are what actually stored.
  Converting binary 01011111 to decimal:
  Decimal value
  \[ = 1\times2^6 + 1\times2^5 + 1\times2^4 + 1\times2^3 + 1\times2^2 + 1\times2^1 + 1 \]
  \[ = 64 + 16 + 8 + 4 + 2 + 1 + 1 = 95 \]
Converting binary 01 011 111 to octal, base 8:
Octal digits: 0 1 2 3 4 5 6 7
Corresponding octal value = 1 3 7

Convert binary 0101 1111 to hexadecimal, base 16:
Hexadecimal digits: 0 1 2 3 4 5 6 7 8 9 A B C D E F
Corresponding hexadecimal value = 5 F

How to treat MIME code (in e-mail, or on web)?

• E-mail transport is ASCII- and line-based.

• Multimedia data, e.g. photo, audio, video, etc., are all byte-based, or in multiples of 8 bits, but not coded in ASCII, and not in terms of lines. They must be converted (encoded).

Reference:
http://www.helpdesk.umd.edu/topics/email/protocols/315/

• Non-text multimedia data must be converted to ASCII, in limited line lengths for e-mail.

• Cannot in a continuous manner simply treat each byte as an ASCII character.
  – There are unprintable characters.
  – There may be no new-line characters, or some of them are in arbitrary positions.

  Problem:
  SMTP has line length limits.
  (76 characters)

• Most use Base64 encoding for byte-based binary data
  – Code every 3 bytes (24 bits) of data into 4 characters (4 bytes long) from a special subset of 64 printable ASCII characters.
    • 6 bits can represent numbers from 0 to 63.
    • We can use 6 bits to index into an array of 64 elements.

• We use 6 bits to index into the special set of 64 characters:
  “ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789+/
  e.g., 0 for A, 1 for B, 63 for .

• Every 3 bytes contains 24 bits, which amount to 4 of 6 bits. Use each of the 6 bits as an index to one of the 64 characters in the special set.

• Thus for every 3 bytes of arbitrary data, it can be represented as 4 printable ASCII characters (from the special set).

  – How to deal with line-length limitation under SMTP?
    • Add additional new-line characters to create short lines artificially.
    • Remove the new-line characters when converting back.
– What if the byte-based data is not in multiple of 24 bits long?
  • Pad with one or two special “=” characters to make the last 24 bits.
  • The special characters are remove when converting back.

– Base64 requires more bytes, approximately 33% extra, to represent original binary data, but results in limited length, printable ASCII lines, good for SMTP.
  • There are other encoding methods, but may be less efficient.

– Base64 encoding example, from http://en.wikipedia.org/wiki/Base64

Original text:
Man is distinguished, not only by his reason, but by this singular passion from other animals, which is a lust of the mind, that by a perseverance of delight in the continued and indefatigable generation of knowledge, exceeds the short vehemence of any carnal pleasure.

Encoded in base64 is as follows:
TWFuIGlzIGRpc3Rpbmd1aXNoZWQsIG5vdCBvbmx5IGJ5IGhpcyByZWFzb24sIGJ1dCBieSB0aGlzIHNpbmd1bGFyIHBhc3Npb24gZnJvbSB0aGUgY29udGludWVkIGFuZCBpbmRlZmF0aWdhYmxlIGdlbmVyYXRpb24gb2Yga25vd2xlZGdlLCBleGNlZWRzIHRoZSBzaG9ydCB2ZWhlbWVuY2UgdGhlIGlvdGx5IGRldmVsdGVyIHdlIGRvY3VtZW50aW9uIG9uIHJlc2VsyW1sIAo=
The coding, in 5 lines, is obtained by treating the original text as a set of binary data.

E-mail via telnet
hutch:/home/shen/public_html/cs312> telnet mail.cs.odu.edu
25
Trying 128.82.4.9...
Connected to mail.cs.odu.edu.
Escape character is "^]".
220 cartero.cs.odu.edu ESMTP
helo kago2.cs.odu.edu
250 cartero.cs.odu.edu Hello hutch.cs.odu.edu [128.82.4.3], pleased to meet you
mail from: fakeSender@cs.odu.edu
250 2.1.0 fakeSender@cs.odu.edu... Sender ok
rcpt to: shen@cs.odu.edu
250 2.1.5 shen@cs.odu.edu... Recipient ok
data
354 Enter mail, end with "." on a line by itself
subject: faked sender
Now, you see that the faked sender still goes through.
.
250 2.0.0 i8TKPsrB007677 Message accepted for delivery quit
221 2.0.0 cartero.cs.odu.edu closing connection
Connection to mail.cs.odu.edu closed by foreign host.
hutch:/home/shen/public_html/cs312>

data
354 Enter mail, end with "." on a line by itself
subject: faked sender
Now, you see that the faked sender still goes through.
.
250 2.0.0 i8TKPsrB007677 Message accepted for delivery quit
221 2.0.0 cartero.cs.odu.edu closing connection
Connection to mail.cs.odu.edu closed by foreign host.
hutch:/home/shen/public_html/cs312>

• The mail received from the previous faked sender

• E-mail via telnet

hutch:/home/shen/public_html/cs312> telnet mail.cs.odu.edu
25
Trying 128.82.4.9...
Connected to mail.cs.odu.edu.
Escape character is "^]".
220 cartero.cs.odu.edu ESMTP
helo kago2.cs.odu.edu
250 cartero.cs.odu.edu Hello hutch.cs.odu.edu [128.82.4.3], pleased to meet you
mail from: fakeSender@cs.odu.edu
250 2.1.0 fakeSender@cs.odu.edu... Sender ok
rcpt to: shen@cs.odu.edu
250 2.1.5 shen@cs.odu.edu... Recipient ok
data
354 Enter mail, end with "." on a line by itself
subject: faked sender
Now, you see that the faked sender still goes through.
.
250 2.0.0 i8TKPsrB007677 Message accepted for delivery quit
221 2.0.0 cartero.cs.odu.edu closing connection
Connection to mail.cs.odu.edu closed by foreign host.
hutch:/home/shen/public_html/cs312>
**IMAP**
- Internet Message Access Protocol (IMAP)
  - An application layer Internet standard protocol used to retrieve e-mail from a remote server to a local client over a TCP/IP connection.
  - Stores messages on the server, can selectively delete messages.


- IMAP 4 has many additional features than POP3
  - Both connect and disconnect modes
  - Multiple clients simultaneous connection to the same mailbox
  - Partial fetch from multi-MIME parts
  - Message state information kept on server
  - Multiple mailbox support
  - Server-side search
  - Extensions to base protocol
  - Not universally supported
    *Previous experience had reliability problem.*

---

**POP3**
- Post Office Protocol version 3 (POP3)
  - An application layer Internet standard protocol used to retrieve e-mail from a remote server to a local client over a TCP/IP connection.
  - Users generally
    - connect
    - retrieve all messages


- Supported by most all ISPs, and is the most popular one.
- store them on the user's PC as new messages
- delete them from the server
- disconnect.
- Messages could be left on the server, but no selective deletion.

---

**The e-mail message format**
- Two major components
  - **Header** - Usually containing:
    - From (Does not have to be the address of the real sender.)
    - To
    - Subject
    - Date
    - Others: Cc, Bcc, Content-Type (usually a MIME type), etc.
  - **Body** - The message itself, usually containing a signature block at the end

- POP3 currently supports
  - several authentication methods to provide varying levels of protection
  - encrypt using Secure Socket Layer (SSL).
Mail User Agent

- GUI mail user agent
  - Freely available
  - Allow choosing between POP3 or IMAP
  - ODU CS mail server does not provide IMAP service
- Very easy to use
- Many additional features
- Example: Netscape Mail
  Discussions in textbook, 1.6

E-mail advantages

- It's free
  Provided that you can get onto the Internet, paid or free.
- It's fast and "assured"
  You get an "instant" response if your e-mail cannot get through.
- Can be sent to many people easily
  You even have a record who the receivers are, "carbon copy" easily.

- Can archive incoming and outgoing e-mails
  Can easily structure the archive, search, etc.
- Can send attachments easily
  Many kinds of electronic files, including photos, documents of many kinds, even music or video files.
- What more would you like to add?
E-mail disadvantages

- Spamming
- Fraud and crime
- Worms
- Attached files with virus
- No response
- Wasting time

Avoid e-mail problems

- Avoid e-mail virus attacks
  - Recognize virus tricks
    - Mostly from attachments
      - Executable files: .exe, .vbs, etc.
      - Disguised file names: appearing to be .txt, .jpg but appended with many blanks then the real extension name
    - Attractive subject line or content
      - Money making opportunities
      - Free offers
      - Friendly warnings

- Computer worm and computer virus
  - Computer virus
    - a self-replicating program
    - spreads by inserting copies of itself into other executable code or documents
    - viruses are one of the several types of malicious software.

- Computer worm
  - A self-replacing program
  - Similar to a computer virus
  - Self-contained, does not need to be part of another program to propagate itself
  - Often designed to exploit the file transmission capabilities found on many computers

Reference:

Free e-mail accounts

- You basically get an e-mail account if you have some account to get onto the Internet.
- You can get free e-mail accounts from various sites. Use a search engine and you can see many such sites.
  - http://www.walla1.com/ 1GB account
  - http://www.yahoo.com/ 250 MB
  - There used to be, and may still have, some sites provide free dial-up with their (limited) free e-mail service.

- Helpful sites
  - http://hoaxbusters.ciac.org/
  - http://www.mail-abuse.com/
  - http://spam.abuse.net/
  - http://www.cauce.org/

Reference:
http://www.cauce.org/
Electronic Mailing List

- (Electronic) Mailing list
  - A collection of names and addresses
  - The group of subscribers to such a list
  - For collaboration on projects, for marketing, etc.
  - Example, Wikipedia mailing list, open to anyone who subscribes: http://meta.wikimedia.org/wiki/Mailing_list

Software example: LISTSERV®

- A system that allows you to create, manage and control electronic mailing lists on a corporate network or on the Internet.
- http://www.lsoft.com/manuals/1.8d/ownerindex.html
- Many features:
  - Customizable Web interface for list administrator and users
  - Virus and spam protection
  - Searchable Web archive
  - Moderated and unmoderated discussion groups
  - Etc., etc.

Newsgroup

- A repository for messages, questions, etc. posted from many users at different locations, a public message board, a discussion group, a
- Arranged into hierarchies, often searchable
- Typically postings have varied predefined life time, e.g. one week for some and two weeks for some others
- Many ISPs provide own News Server

Posting may be moderated or unmoderated

- Example, Google Groups, has archives:
  - http://groups.google.com/
  - Alt. User-created, arbitrary, subgroups
  - Biz. Business products, services, reviews...
  - Comp. Hardware, software, consumer info....
  - Humanities. Fine art, literature, philosophy...
  - ..... To post in a Google Group (need to have a Google account):
  - http://groups.google.com/googlegroups/posting_faq.html

Netiquette

- The conventions of politeness recognized in e-mailing, web boards, etc.
  - Think before you post (or e-mail to an individual or a list. We use these terms interchangeably in here.)
  - Use a sensible subject line.
  - Quote the minimum necessary, and be explicit and clear.
  - Re-read and edit your message, use a spell checker.

- Be concise and to the point. Make short lines if your software does not do automatic wrapping.
- Use short paragraphs and double spaces between them.
- Use point form when you can, but use hyphens - or asterisks * as bullets, not real bullets.
- When asking a question, do enough homework and give details.
- Send or post only compressed photos in general.
– Remember that many e-mail clients cannot handle fancy formats.
– Avoid jargon terms or abbreviations.
– E-mail salutations: Differentiate business from personal
– Signatures: In general cases, give full name and return e-mail address. (It maybe not easy to figure out the return e-mail address.)

– Avoid attachment if appropriate.
– Reply only to those you need to – not just to all.
– Remember that e-mails are not private.
– Flames: Best to ignore it and try not to cause one. If you believe someone has violated netiquette, send a message by private e-mail; do not post a follow-up to the news.
– Unsubscribe appropriately, re-subscribe when need to.
– Use junk mail filter.

Instant Messenger

• Allows you to communicate with others interactively over the Internet

• Example: AOL Instant Messenger
  – Can add and structure screen names
  – Can communicate with any persons who are logged in

Internet Telephony

• Use of IP data connections to exchange voice and fax data (that have traditionally been carried over the public switched telephone network)

• Based on voice over IP (VOIP) to be used over the Internet
  – being deployed on internal corporate networks, and, via the Internet

• Internet Telephony Service Providers (ITSPs)
  – Companies providing these services, including telephone companies, cable TV companies and ISPs
  – For low cost (and lower quality) international calls
  Not using dedicated phone lines, but using packet switching instead

Anybody has firsthand experiences with it?

• Example: the WebPhone
  Home page:
hhttp://www.webphone.com/index.html
  Tutorial:
hhttp://www.webphone.com/tutorial.html
  Better voice quality with broadband connection.

  Not “cheaper” nor “more convenient” than, e.g. masterbell calling card:
hhttp://www.masterbell.com/
Conclusion

- E-mails started early, before the Internet
- Non-ASCII data needs MIME encoding
- Encoding increases data file size
- Should follow netiquette
- Watch for computer virus
- Many new communication technologies over the Internet