Preparation and Comparison of Nicknames
draft-saintandre-precis-nickname-00

Abstract

   This document describes how to prepare and compare Unicode strings
   representing nicknames, primarily as used within textual chatrooms.
   This profile is intended to be used by chatroom technologies based on
   both the Extensible Messaging and Presence Protocol (XMPP) and the
   Message Session Relay Protocol (MSRP).

Status of this Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF). Note that other groups may also distribute
   working documents as Internet-Drafts. The list of current Internet-
   Drafts is at http://datatracker.ietf.org/drafts/current/.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time. It is inappropriate to use Internet-Drafts as reference
   material or to cite them other than as "work in progress."

   This Internet-Draft will expire on September 6, 2012.

Copyright Notice

Table of Contents

1. Introduction

1.1. Overview

   Technologies for textual chatrooms customarily enable participants to
   specify a nickname for use in the room; e.g., this is true of
   Internet Relay Chat [RFC2811], Multi-User Chat (MUC) based on the
   Extensible Messaging and Presence Protocol (XMPP) [XEP-0045], and
   multi-party chat based on the Message Session Relay Protocol (MSRP)
   [I-D.ietf-simple-chat]. Recent chatroom technologies also allow
   internationalized nicknames because they support characters from the
   outside the ASCII range, typically by means of the Unicode character
   set [UNICODE]. Although such nicknames are often used primarily for
   display purposes, they are sometimes used for programmatic purposes
   as well (e.g., kicking users or avoiding nickname conflicts).

   To increase the likelihood that nickname input and comparison will
   work in ways that make sense for typical users throughout the world,
   this document defines rules for preparing and comparing
   internationalized nicknames.

1.2. Terminology

   Many important terms used in this document are defined in
   [I-D.ietf-precis-framework], [RFC6365], and [UNICODE]. Relevant XMPP
   terms are defined in [RFC6120] and [XEP-0045], and relevant MSRP
   terms in [RFC4975] and [I-D.ietf-simple-chat].

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
   "OPTIONAL" in this document are to be interpreted as described in RFC
   2119 [RFC2119].

2. Rules

   A nickname MUST NOT be zero bytes in length and MUST NOT be more than
   1023 bytes in length (the latter restriction is derived from the
   length restriction on XMPP resourceparts, see [RFC6122]). This rule
   is to be enforced after any mapping or normalization of code points.

   A nickname MUST consist only of Unicode code points that conform to
   the "FreeClass" base string class defined in
   [I-D.ietf-precis-framework].

   For preparation purposes (e.g., when a chatroom client generates a
   nickname from user input for inclusion as a nickname protocol
   element), an application MUST only ensure that the string conforms to

the "FreeClass" base string class defined in
[I-D.ietf-precis-framework]; however, it MAY also perform the mapping
and normalization operations specified below for comparison.

For comparison purposes (e.g., when a chatroom server determines if
two nicknames match during the authorization process), an application
MUST treat a nickname as follows, where the operations specified MUST
be completed in the order shown:

1. Non-ASCII space characters from the "N" category defined under
    Section 6.14 of [I-D.ietf-precis-framework] MUST be mapped to
    SPACE [U+0020].

2. Uppercase and titlecase characters MUST be mapped to their
    lowercase equivalents. In applications that prohibit matching
    nicknames, this rule helps to reduce the possibility of confusion
    by ensuring that nicknames differing only by case (e.g.,
    "stpeter" vs. "StPeter") would not be allowed in a room at the
    same time.

3. All characters MUST be mapped using Unicode Normalization Form KC
    (NFKC). Because NFKC is more "aggressive" in finding matches
    than other normalization forms (in the language of Unicode, it
    performs both canonical and compatibility decomposition before
    recomposing code points), this rule helps to reduce the
    possibility of confusion by increasing the number of characters
    that would match (e.g., ROMAN NUMERAL FOUR [U+2163] would match
    the combination of LATIN CAPITAL LETTER I [U+0049] and LATIN
    CAPITAL LETTER V [U+0056]).

For both preparation and comparision, the "Bidi Rule" provided in
[RFC5893] applies to the directionality of a nickname.

3. Security Considerations

3.1. Reuse of PRECIS

The security considerations described in [I-D.ietf-precis-framework]
apply to the "FreeClass" base string class used in this document for
nicknames, respectively.

3.2. Reuse of Unicode

The security considerations described in [UTR39] apply to the use of
Unicode characters in nicknames.

3.3. Visually Similar Characters

[I-D.ietf-precis-framework] describes some of the security
considerations related to visually similar characters, also called
"confusable characters" or "confusables".

Although the mapping rules under Section 2 are designed in part to
reduce the possibility of confusion about nicknames, this document
does not yet provide more detailed recommendations regarding the
handling of visually similar characters, such as those in [UTR39].
However, a future version of this document might provide such
recommendations.

4. IANA Considerations

The IANA shall add an entry to the PRECIS Usage Registry for reuse of
the PRECIS FreeClass for preparation and comparision of nicknames, as
follows:

Application Protocol: MSRP and XMPP.
Base Class: FreeClass
Subclassing: No.
Directionality: The "Bidi Rule" defined in RFC 5893 applies.
Casemapping: None.
Normalization: NFC.
Specification: RFC XXXX.

5. References

5.1. Normative References

[I-D.ietf-precis-framework]
            Blanchet, M. and P. Saint-Andre, "Precis Framework:
            Handling Internationalized Strings in Protocols",
            draft-ietf-precis-framework-01 (work in progress),
            October 2011.

[I-D.ietf-simple-chat]
            Niemi, A., Garcia, M., and G. Sandbakken, "Multi-party
            Chat Using the Message Session Relay Protocol (MSRP)",
            draft-ietf-simple-chat-14 (work in progress), March 2012.

[RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
            Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5893]   Alvestrand, H. and C. Karp, "Right-to-Left Scripts for

            Internationalized Domain Names for Applications (IDNA)",
            RFC 5893, August 2010.

   [RFC6122] Saint-Andre, P., "Extensible Messaging and Presence
              Protocol (XMPP): Address Format", RFC 6122, March 2011.

   [UNICODE] The Unicode Consortium, "The Unicode Standard, Version
              6.1", 2012,
              <http://www.unicode.org/versions/Unicode6.1.0/>.

   [UTR39]   The Unicode Consortium, "Unicode Technical Report #39:
              Unicode Security Mechanisms", August 2010,
              <http://unicode.org/reports/tr39/>.

   [XEP-0045]
              Saint-Andre, P., "Multi-User Chat", XSF XEP 0045,
              February 2012.

5.2. Informative References

   [RFC2811] Kalt, C., "Internet Relay Chat: Channel Management",
              RFC 2811, April 2000.

   [RFC4975] Campbell, B., Mahy, R., and C. Jennings, "The Message
              Session Relay Protocol (MSRP)", RFC 4975, September 2007.

   [RFC6120] Saint-Andre, P., "Extensible Messaging and Presence
              Protocol (XMPP): Core", RFC 6120, March 2011.

   [RFC6365] Hoffman, P. and J. Klensin, "Terminology Used in
              Internationalization in the IETF", BCP 166, RFC 6365,
              September 2011.

Author's Address

   Peter Saint-Andre
   Cisco Systems, Inc.
   1899 Wynkoop Street, Suite 600
   Denver, CO 80202
   USA

   Phone: +1-303-308-3282
   Email: psaintan@cisco.com