

Design of registration and validation algorithm of member's personal data

Solomia Fedushko, Yuriy Syerov

Social Communications and Information Activities Department, Lviv Polytechnic National University

Article Info

Article history:

Received Mar 19th, 2013

Revised May 20th, 2013

Accepted June 30th, 2013

Keyword:

Algorithm

Data verification

Personal profile

Web-community member

ABSTRACT

This paper deals with the methods of verification of web-forum members' personal data, which are implemented in the algorithm of registration and validation of personal data on web-community member. The problem of personal data verifying and method of registration in an online community is investigated. These issues have the greatest influence on efficiency rise of virtual communities functioning and the level of data authenticity in web-members' personal profiles. Solution of these problems is possible by using computer-linguistic analysis of web-members' posts.

Copyright © 2013 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Fedushko Solomia,
Social Communications and Information Activities Department,
Institute of Humanities and Social Sciences,
Lviv Polytechnic National University,
12 Bandery Street, Lviv, 79013 Ukraine
Email: felomia@gmail.com

1. INTRODUCTION

Nowadays, the necessity of verifying of web-community members' personal data is the topical issue. Thus, this linguistic research of difference styles of writing posts is the basis for developing effective methods of verifying of personal data in users account.

Web-communities accumulated a huge database of contacts and profiles, which contain a lot of information about the person. The auditorium of web-communities is a large number of people regardless of age, gender, occupation, education, ethnicity, social status etc., who in order to register need to fill in a form with their personal data.

Practical value of the research is determined by the development of an algorithm of registration and validation of personal data on web-community member. The algorithm of registration and validation of personal data is formed on the basis of specialized software which has been implemented in work of *Lviv.Ridne Misto* web-forum. The developed approaches to registration and validation of personal data on web-community member have been analyzed.

The algorithm of registration and validation of account data on web-community member helps to classified web-community members by level of authenticity of the information in members' profiles.

Thus, computer-linguistic analysis can be of essential help to administrators and moderators to monitor and validate personal data provided by users of the web-forum.

2. REGISTRATION METHODS OF WEB-FORUM MEMBERS

In this paper issue of personal data verifying and method of registration in an online community is invastigated. These issues have the greatest influence on efficiency rise of virtual communities functioning and the level of data authenticity in web-members' personal profiles. Solution of these problems is possible

by using computer-linguistic analysis of web-members' posts. Thus, the actual problem of modern web is "information noise" [3] - a huge amount of inexact, incomplete and superficial information, which Internet users created on the Web [4].

Registration methods of web-forum members [9] are divided into three groups:

- registration without verifying (anonymous);
- registration with verifying;
- registration by invitation.

2.1. Registration without verifying.

These are the methods of participation in web-community. This is the easiest method of participation in virtual communities. The method of participation without verifying provides two variants: member voluntary gives his name (nickname) or participates in communications without providing the name (anonymous participation). The disadvantage of this verifying method is the lowest level of protection from intruders; the advantage is the highest probability that users will create the message.

2.2. Registration with verifying.

This method of registration requires from a person who is registered to provide certain information, which is foreseen of the web-forum regulations in the simplest case, verifying can be fully automatic and requires only correct email address. The registration is partially automated and requires web-community moderator and/or administrator participation, in the complicated case. At this rate the verifying process has different levels of difficulty.

2.3. Registration by invitation.

This method of registration is not very distributed and used in specific close-type communities. Verifying of the member had occurred before the process of registration.

Owners of web-forum choose one of the methods of member's registration [9]. It depends on determined script of development web-forum. Today the most comprehensive method is the registration with verifying. This method has different levels of difficulty depending on the wishes of the owners of web-forum.

3. ALGORITHM OF REGISTRATION AND VALIDATION OF PERSONAL DATA

The method of personal data validation of the maximum amount of information about potential web-forum member is developed and basic foundation for computer-linguistic method of web-forum members' information track verifying [7], which is based on linguistic analysis of web-community content is devised.

Algorithm of premoderated registration for filtration of undesirable web-forum members is destined. The idea of the algorithm contains the maximum verifying and information analysis of the members' data, which made the enquiry about registration.

Algorithm of registration and validation of account data on web-community member (Figure 1):

1. Email addresses validation.

The email address is incorrect. The process of registration is stopped because administrator can not send a password.

2. Internet name validation according to the web-community regulations.

The email address is correct, than the Internet name verifying takes place:

- in accordance with the naming rules of web-community;
- in the presence of already registered members with this Internet name.

3. Computer-linguistic verifying of the Internet name.

The Internet name to web-community regulations is conforms and it is free. The computer-linguistic analysis is occurred [7]. Computer-linguistic analysis based on gender and age verifying of Internet name. In this analysis the categories of Internet name selection are used.

Internet name. In the context of information technology, an Internet name (nickname or technically a nick is a common synonym for a screen name or handle) is a term originally used to identify a person in a system for synchronous conferencing. In computer networks it has become a common practice for every person to also have one or more nicknames for the purposes of anonymity, to avoid ambiguity or simply because the natural name or technical address would be too long to type or take too much space on the screen.

Members of web-communities take their Internet names into consideration because they are person's on-line alter ego, and they help users to emphasize their online identity. Sometimes Internet names can say more about people, their hobbies and temper than the personal information in the users account.

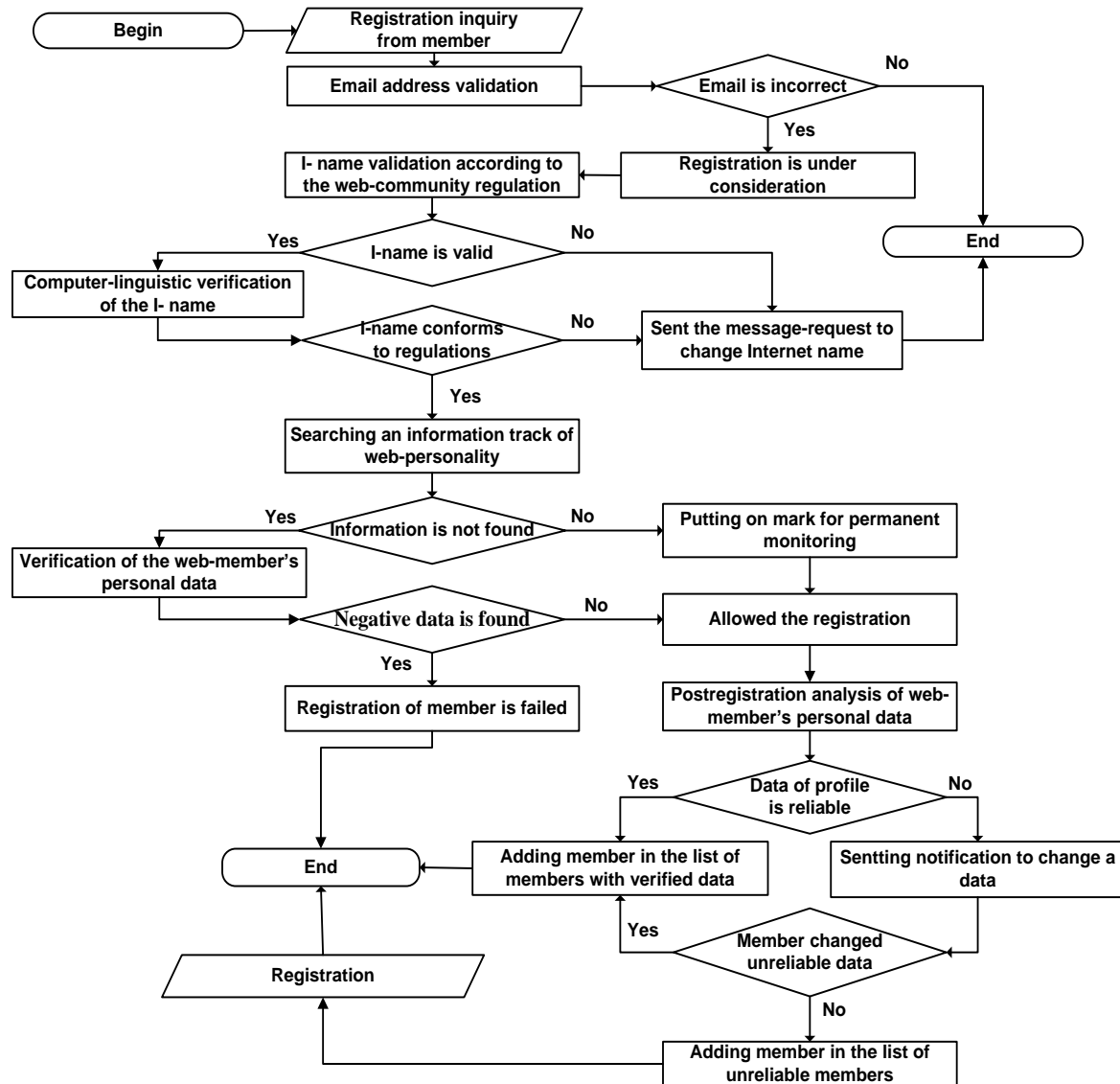


Figure 1. Algorithm of registration and validation of personal data on web-community member

Internet names become an essential part of a child's explanation of themselves and the world they inhabit and a natural focus of early childhood questioning. Names help children and adolescents identify 'who they are', understand their 'self' and have an effect on their approach to social interactions [1]. In short, names not only help to distinguish and describe a child's internal and external world, but also the ways in which they are perceived by other children.

Nicknames also affect social interaction. Positive nicknames provide ways to fit in among peer groups [10], while negative ones may affect self-image or peer perceptions. Nicknames are often given by the people that surround us, but a self-chosen nickname may augment a person's self-esteem. Virtual persons are constructed as online users choose their names and nicknames [4, 6, 7]. Not only is a username needed to log onto the Internet, but in all forms of online interactions: emailing; chatting; posting; even logging into web-sites. Therefore, names provide the primary form of online identity.

One of the goals of this study is to understand what types of names authors choose when creating and maintaining their web-forum personal accounts. Author names were classified into two categories: children and adult online names.

In this research we defined gender and age conformity in online name choice by members of web-communities. The results of my research are: adult users tended to use their real names more often forms than children. Children do not choose their real names, instead they choose: boys tend to choose nicknames, nonsensical names, names related to fantasy, mythology, films and tales, instruments or act of violence, cars, sports and music and pop culture while girls tend to choose plants, animals, gentle and pet names, self-related names, names related to music and pop culture and names related to natural phenomenon.

There were gender differences for online name choice – males would choose realistic names more often than females. Females tend to choose names of natural phenomenon, plants, animals, names related to literature, films, fairytales, mythology, history or famous people, temper, mood and moral description names and using a figure of speech in creating names while males to choose names which are related to political convictions, profession-related names, names of location and names related to literature, films, fairytales, mythology, history or famous people. In this research web-communities were analyzed to understand what types of Internet names teenage boys and girls, men and women choose.

When the Internet name was already used by another member or it did not conform to web-community regulations and computer-linguistic verifying of Internet name is unsuccessful, the message-request to change his Internet name to the applicant is sent.

4. Verifying of the web-member's personal information.

If the Internet name of potential member is verified then the verifying of other member's personal data (it requires participation of administrator) is conducted.

In particular, the following data is verified:

- e-mail address and all additional information is validated in accordance with the Internet name and authenticity by search engine means;
- authenticity of the geographical position indication by tools of IP-addresses determining is verified [9];
- information track of web-personality [11], [8] of the potential web-forum member is verified by search engines.

When information about the member is not found, registration is allowed and the mark is put on to the effect that to permanently monitored the web-forum member's communicative behavior and analysed member's personal information verifying.

The given information is not compromised the member and information track of web-personality is not contain negative information. The member the approval to join to the community is received. The computer-linguistic analysis of the personal data of member's profile is conducted [**Error! Reference source not found.**, p.89]. Otherwise, the registration of this member is failed.

5. Postregistration computer-linguistic analysis of avatars, userbars, biographies, signatures and status of web-members.

There is verifying of the information, which by web-forum members in their personal accounts is represented. According to the result of this analysis, it is possible to identify certain characteristics of member (age, gender, biography, nationality, education etc.). If the data of member's profile is unreliable, then the notification to change a personal data to member is sent.

Request was ignored and member did not change the personal data in profile (age, gender, occupation, biography, etc. data.), then the member in the list of unreliable members is added. All actions of unreliable members are monitored. The reliable data in members' profiles is given or it is changed at the instance of administration. The members in the list of member with verified data are added and all further activities of the members are monitored.

Method of registration and personal data validation are realized in the algorithm of registration and personal data validation on web-community member. The main idea of the algorithm is the maximum verifying and analyse of web-community members' information that the inquiry about registration is made. Postregistration computer-linguistic verifying of personal data in web-community members' profiles and their information trace is analyzed by means of algorithm of registration and validation of personal data on web-community member.

This algorithm increases number of members with complete (reliably) information and gives the possibility to classify them, to improve the communicative behavior of members and to diminish pseudomembers from web-community. The algorithm the web-forum management helps to increase the web-forum efficiency. A question of urgent importance in the web-forum management and moderation is the developing of a new approach to data verifying which gives web-members when they are resisted.

4. CLASSIFICATION OF WEB-FORUM MEMBERS' ACCOUNTS INTO LEVEL OF RELIABILITY DATA

A conclusion section is not required. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

The algorithm of registration and validation of account data on web-community member helps to classified web-community members by level of authenticity of the information in members' profiles.

Analysis of registration and personal data verifying of members, which will occur according to algorithm of registration and validation of account data on Web-community member allows to detected level of authenticity of data in web-members' personal profiles and to separated members with reliable and unreliable data.

Classification of Web-forum members' accounts into level of data reliability:

- a. account with reliable information;
- b. account data under suspicion;
- c. account with inadequate information (pseudo user account).

4.1. Account with reliable information.

Personal information which gives member in his account is reliable or member changed it at the request of web-community administration which is sent during computer-linguistic analysis. The communicative behavior of the member conforms to all web-community regulations.

4.2. Account data under suspicion.

Not all personal data of web-members are reliable. The member partially or completely ignored the request to change the incorrect information. All data changes and communicative behavior are monitored by web-community administration.

4.3. Account with inadequate information (Pseudo user account).

The data in personal profile is completely unauthentic. The Web-forum member contravenes the web-community regulations. The member profile is blocked / deleted or access to the web-community is banned. The result of research is classification of web-forum members' accounts, which helps administrator to arranged web-forum members' accounts and promptly monitored the personal data in web-forum members' accounts.

5. CONCLUSIONS

Computer-linguistic analysis of web-communities members' information track is the basic way to develop methods of verifying personal data of web-forum members' accounts based on analysis of gender and age similarities and differences in online identity and language use of web-forum members. Computer-linguistic analysis helps administrators and moderators to monitor and validate personal data provided by users of the web-forum.

The issue of personal data verifying and method of registration in an online community is investigated. web-forum members' registration methods are investigated and divided into types. The method of personal data validation of the maximum amount of information about potential web-forum member is developed and basic foundation for computer-linguistic method of web-forum members' information track verification, which is based on linguistic analysis of web-community content, is devised. A question of urgent importance in the web-forum management and moderation is the development of a new approach to data verification which gives web-members when they are resisted.

Analysis of registration and personal data verification of members which take place according to algorithm of registration and validation of account data on web-community member allows to detect authenticity level of data in web-members' personal profiles.

The result of the research is classification of web-forum members' accounts which can help an administrator to arrange web-forum members' accounts.

The paper presents a new approach to developing computer-linguistic method of web-member identification by some characteristics and identifying web-forums members by means of computer-linguistic analysis of web-communities members' information tracks.

REFERENCES

- [1] Calvert S., Jordan, A. and Cocking, R. Children in the digital age: Influences of electronic media on development. Westport, CT: Praeger, 2002, pp. 3-33
- [2] Calvert S., Jordan, A., Cocking, R. Identity construction on the internet. Children in the Digital Age: Influences of Electronic Media on Development: Praeger, Westport, CT, 2002, pp. 57-70.
- [3] Carr N. (October 3, 2005) The amorality of Web 2.0. Rough type. Available: http://www.roughtype.com/archives/2005/10/the_amorality_o.php
- [4] Croll A. and S. Power, Complete Web Monitoring: Watching Your Visitors, Performance, Communities, and Competitors, O'Reilly Media, 2009, pp.672.

- [5] Fedushko S. 'Disclosure of Web-members Personal Information in Internet'. Conference "Modern information technologies in economics, management and education", Lviv, 2010, pp.163-165.
- [6] Greenfield P., Subrahmanyam K. "Online discourse in a teen chatroom: New codes and new modes of coherence in a visual medium", Journal of Applied Developmental Psychology, 2003, ch.24, pp. 713-738.
- [7] Peleschyshyn A., Fedushko S. "Gender similarities and differences in online identity and Internet communication". Proceedings of the International Conference of Computer Science and Information Technologies «CSIT-2010», Lviv, 2010, pp.195-198.
- [8] Peleschyshyn A. and Yu. Syerov, "Typical Ways of Web Communities Development", Proceedings of the International Conference of Computer Science and Information Technologies «CSIT-2006», Lviv, 2006, pp.56-58.
- [9] Peleschyshyn, A., Syerov, Yu. and Fedushko, S. 'Developing algorithm of registration and validation of personal data on web-community member'. Journal of the Lviv Polytechnic National University: Computer Science and Information Technology, Lviv, (ed.) NU "LP", №686, 2010, pp. 238-244.
- [10] Reid E. "Electropolis: Communication and community on the Internet Relay Chat". Department of History, University of Melbourne, Melbourne, Australia, 1991.
- [11] Syerov, Yu. and Kravets R. 'Software for determination of behavior patterns of web-forum members'. International Journal of Computing "Computing", Ternopil, Vol. 8(2), 2009, pp.119-125.

BIOGRAPHY OF AUTHORS



Solomia Fedushko

Master of science in applied linguistics.

At the present moment: PhD student and assistant at the Social Communications and Information Activities Department, Lviv Polytechnic National University.

Scientific interests: internet linguistics, social communications in Web, web-communities, web-community content analysis.



Yuriy Syerov

Science and Engineering

Master of computer sciences, Lviv Polytechnic National University.

Assistant at department of Information systems and networks (2004-2010). Received PhD degree in technical sciences from Lviv Polytechnic National University, in 2010.

At the present moment: Associated Professor at the Social Communications and Information Activities Department.

Scientific interests: web-communities, organization of effective web-communities, web-mining technologies.