

SUSTAV ZA AUTOMATSKU DETEKCIJU LAŽNIH PORUKA

Marin Vuković

Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu

marin.vukovic@fer.hr

Krešimir Pripuzić

Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu

kresimir.pripuzic@fer.hr

Sažetak

Rad se bavi problemom definicije i uočavanja lažnih poruka (hoax). Predložen je sustav za uočavanje, kategorizaciju i rangiranje lažnih poruka. Opisani su postupci pretprocesiranja sadržaja poruka, razlikovanja običnih i sumnjivih poruka te dodatne analize i kategorizacije poruka za koje se sumnja da su lažne.

1. Poruke neistinitog sadržaja (hoaxes)

Razvoj Interneta i usluga poput elektroničke pošte i trenutnog poručivanja (*instant messaging*), uz mnoge prednosti, stvorio je i plodno tlo za različite malverzacije, krađe i slične ilegalne aktivnosti. Jedan od načina izvođenja takvih aktivnosti je i razaslanje poruka neistinitog sadržaja koje nazivamo *hoaxovima* [1]. Već se iz samog naziva termina može zaključiti kako je glavna ideja kreatora takve poruke prevariti i nagovoriti primatelja na željenu aktivnost, spriječiti unaprijed dogovorene aktivnosti, odnosno uvjeriti ih u činjenice koje nisu stvarne [2].

Takve poruke, promatrajući s psihološkog aspekta, kreirane su tako da iskorištavaju ljudsku potrebu da se pomogne drugom ili s druge strane, pohlepu koja tjera čovjeka da pridobije lažno ponuđen novac ili neko drugo materijalno sredstvo. Poruke neistinitog sadržaja same po sebi ne predstavljaju izravnu opasnost po računalo, dlanovnik ili slično, već svojim sadržajem mogu natjerati primatelja da štetno djeluje na računalni sustav, vlastitu i reputaciju tvrtke te mogu uzrokovati različite financijske gubitke.

Krađe osobnih informacija, kao što su korisnička imena, lozinke, brojevi kartica i slično, prema autorima [1] mogu rezultirati visokim pojedinim štetama. Navedeno je dodatni motiv da se danas većina država putem različitih organizacija aktivno bori da zaustavi stvaranje i širenje takvih poruka.

2. Sustav za detekciju poruka neistinitog sadržaja

Sustav za detekciju poruka neistinitog sadržaja objedinjuje postupke pretprocesiranja poruka, uočavanja sumnjivih poruka iz skupa poruka te detaljne analize i kategorizaciju poruka za koje se ustvrdilo da su zaista lažne.

Postupak pretprocesiranja poruka svodi se na uočavanje riječi i fraza karakterističnih za lažne poruke te njihovo prevođenje u normalizirani oblik. Navedene riječi i fraze formiraju skup na temelju kojeg se izvodi uočavanje eventualnih novih lažnih poruka tijekom rada sustava.

Uočavanje lažnih poruka provodi se korištenjem umjetne neuronske mreže, učene na temelju normaliziranih fraza i riječi karakterističnih za poruke lažnog sadržaja [3]. U tom smislu, rezultat ovog procesa jest vjerojatnost da je ulazna poruka lažna. Prema vjerojatnostima je moguće doći do tri razine sigurnosti: da je poruka istinita, sumnjiva ili lažna.

U trećoj fazi se utvrđene sumnjive i lažne poruke kategoriziraju s pomoću samoorganizirajuće mreže [4], kako bi se utvrdilo o kojem se tipu lažne poruke radi i tako krajnjeg korisnika dodatno upozorilo o mogućoj prijetnji.

3. Literatura

[1] Hernandez, J.C., Hernandez, C.J., Sierra, J.M., Ribagorda, A.: A First Step towards Automatic Hoax Detection. In: Proceedings of the International 36th Annual Carnahan Conference on Security Technology, IEEE, Piscataway NJ, pp. 102--114 (2002)

[2] Richardson, R.: 2008 CSI Computer Crime & Security Survey. Computer Security Institute, San Francisco, CA, URL: <http://www.gocsi.com/> (2008)

[3] D. Jevtić, Ž. Car, M. Vuković: Location Name Extraction for User Created Digital Content Services // Lecture Notes in Computer Science/Lecture Notes in Artificial Intelligence LNAI 4692. 1 (2007) ; 623-630

[4] M. Vuković, K. Pripuzić, H. Belani: An Intelligent Automatic Hoax Detection System // Lecture Notes in Computer Science/Knowledge-Based and Intelligent Information and Engineering Systems 5711/2009. 1 (2009) ; 318-325