Neuronska kriptografija kao odgovor na napad kvantnim računalom

Stjepan Picek i Marin Golub
Zavod za elektroniku, mikroelektroniku, računalne i inteligentne sustave
Fakultet elektrotehnike i računarstva
Unska 3, Zagreb, Hrvatska
E-mail: stjepan@computer.org i marin.golub@fer.hr

Sažetak - Kriptografija javnog ključa daje jedan od mogućih odgovora na problem razmjene ključeva između dviju strana. Uobičajeno je da se u kriptografiji javnog ključa sigurnost sustava temelji na "teškim" matematičkim problemima. Najpoznatiji kriptosustav temeljen na tim principima je RSA kriptosustav osmišljen pred više od 30 godina. Mnogo su manje poznate, a još i manje korištene metode čija sigurnost ne počiva na teoriji brojeva. Glavna prednost takvih metoda je što nisu ranjive na napad kvantnim računalom. U takve metode spada i neuronska kriptografija. Neuronska kriptografija se temelji na sinkronizaciji dvaju umjetnih neuronskih mreža. Sigurnost takvom sustavu daje činjenica da je to NP težak problem. Svrha ovog članka je prikazati kako neuronske mreže mogu stvoriti zajednički ključ razmjenom bitova preko javnog kanala i međusobnim učenjem te razmotriti prednosti koje takav protokol ima nad drugim, češće korištenim protokolima. Također, razmatramo slabosti koje čine neuronsku kriptografiju podložnom nekim vrstama napada.

I. UVOD

Moderni kriptografski sustavi se mogu podijeliti na simetrične kriptosustave i kriptosustave s javnim ključem. Cilj svakog kriptografskog sustava je sigurna komunikacija. Preduvjet sigurnosti simetričnih kriptosustava je tajnost ključa. Taj preduvjet predstavlja i glavni nedostatak simetrične kriptografije. Da bi dvije strane, pošiljatelj Ana, i primatelj Branko razmijenile poruku moraju prije toga razmijeniti tajni ključ. Tu se postavlja pitanje ako Ana i Branko imaju na raspolaganju kanal kojim mogu sigurno razmijeniti ključ zašto ne bi tim kanalom razmijenili i poruku. 1976. godine dano je jedno moguće rješenje problema razmjene ključeva u vidu kriptografije javnog ključa. Okosnica takve kriptografije je konstrukcija kriptosustava kod kojih poznavanje funkcije kriptiranja ne osigurava izračunavanje funkcije dekriptiranja. Sigurnost takvih sustava najčešće počiva na „teškim" matematičkim problemima. Glavni nedostatak takvih kriptosustva je njihova sporost. Hibridni kriptosustavi, koji predstavljaju kombinaciju prethodno navedena dva tipa kriptosustava pokušavaju iskoristiti najbolje od oba svijeta. Kriptografski sustav s javnim ključem služi za generiranje ključa a simetrični kriptosustav za šifriranje poruka prethodno generiranim ključem. [1][6]

U zadnjih dvadesetak godina razvile su se nove vrste kriptografija od kojih neke pokazuju visoku sigurnost i učinkovitost, a temeljene su na konceptima poznatima i preko 50 godina. Takve vrste kriptografije nisu temeljene na teoriji brojeva već iskorištavaju određene prirodne fenomene. U toj kategoriji svakako vrijedi spomenuti kvantnu kriptografiju [10], i neuronsku kriptografiju.

Neuronska kriptografija predstavlja granu kriptografije posvećenu analizi primjena stohastičkih algoritama, posebice umjetnih neuronskih mreža u kriptologiji. Za sada nema praktičnih primjena neuronske kriptografije, ali zahvaljujući prednostima koje će biti razmatrane u ostatku članka očekuje se da će se to brzo promijeniti.

Neuronska kriptografija predstavlja veoma mladu granu kriptografije; prvi je puta spomenuta i upotrijebljena 1995. godine u magistarskom radu Sebastiana Dourlensa. On je iskoristio neuronsku kriptografiju u kriptoanalizi DES algoritma na način da je umjetnu neuronsku mrežu učio kako izračunati inverz S-tablica. [12]

U drugom poglavlju predstavljeni su koncepti potrebni za dublje razumijevanje tematike, treće poglavlje obrađuje neuronsku razmjenu ključeva, u četvrtom poglavlju su predstavljani poznati napadi na sustave štićene neuronskom kriptografijom.

II. NEURONSKE MREŽE I SINKRONIZACIJA

A. Neuronske mreže

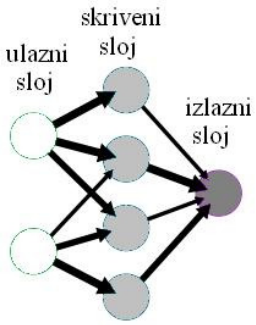
Neuronske mreže se mogu podijeliti na biološke (prirodne) neuronske mreže i umjetne neuronske mreže. Definicija umjetne neuronske mreže koju su koristili Alexander i Morton 1990. godine kaže da je umjetna neuronska mreža masivno paralelni distribuirani procesor koji je dobar za pamćenje iskustvenog znanja.[9] U nastavku rada umjetne neuronske mreže će biti zvane neuronske mreže. Prema topologiji mreža koja se temelji na načinu kojim su neuroni međusobno povezani neuronske mreže se mogu podijeliti na:

- mreže bez povratne veze,
- i mreže sa povratnom vezom.

Nadalje, neuronske mreže bez povratne veze se mogu podijeliti na:

- jednoslojne mreže,
- i višeslojne mreže.

Na slici 1. je prikazana jednostavna neuronska mreža bez povratne veze.



Slika 1. Jednostavna neuronska mreža bez povratne veze

B. Perceptron

Perceptron predstavlja najjednostavniju neuronsku mrežu bez povratne veze za klasifikaciju uzoraka koji su linearno separabilni (leže na suprotnim stranama hiperravnine).

Formalna definicija perceptrona može biti izražena kao:

Perceptron je binarni klasifikator koji preslikava realni ulazni vektor x u izlaznu vrijednost $f(x)$.

$$f(x) = \begin{cases} 1, \text{ako je } w * x + b > 0 \\ 0, \text{inace} \end{cases} \quad (1)$$

gdje vrijednost w predstavlja vektor realnih vrijednosti težina, a vrijednost b predstavlja konstantu. [12]

C. Učenje neuronskih mreža

Neuronska mreža mora biti konfigurirana tako da skup ulaznih vrijednosti tvori željeni izlazni skup. Postoje razne metode za postavljanje snage veza između neurona. Jedan način je učenje gdje se eksplicitno postavljaju težina koristeći *a priori* znanje. Drugi način je učenje neuronske mreže o okolini kroz iterativni proces podešavanja sinaptičkih težina i pragova. Može se dati definicija učenja u kontekstu neuronskih mreža: učenje je proces kojim se slobodni parametri neuronske mreže adaptiraju kroz kontinuirani proces stimulacije od okoline u kojoj se mreža nalazi [9]. Postoje tri glavne paradigme učenja neuronskih mreža:

- učenje pod nadzorom,
- učenje bez nadzora i
- učenje podrškom.

Za detaljnije informacije o neuronskim mrežama preporučamo pogledati dodatnu literaturu [5][8][9].

D. Sinkronizacija i neuronska sinkronizacija

Sinkronizacija predstavlja koordinaciju događaja koji djeluju usuglašeno unutar nekog sustava. Sinkronizacija kao fenomen je opažena u mnogim fizikalnim ili biološkim sustavima. Prvo je primjećena kod oscilatora sa slabom spregom, koji su razvili konstantni odnos faza jedan prema drugome. Po postizanju potpune sinkronizacije, može se opaziti da sustavi posjeduju i jednaku dinamiku.

U zadnjih nekoliko godina je otkriveno i da se neuronske mreže mogu sinkronizirati. Na izlazne vrijednosti neuronskih mreža utječu osim ulaznih vektora vrijednosti i vektori težina w. Vrijednosti vektora težina w se mogu mijenjati na temelju dva osnovna algoritma. Kod *batch* algoritma sve vrijednosti su dostupne u istom trenutku te se tada računa optimalni vektor težina. Kod *online* algoritma koristi se samo jedan vrijednost u svakom vremenskom koraku te time potrebne vrijednosti mogu biti kreirane drugom neuronskom mrežom.

Koncept sinkronizacije je doveo i do neuronskih mreža bez povratne veze koje se sinkroniziraju obostranim učenjem. Takve mreže primaju zajedničke ulazne vektore te uče temeljem izlaznih vrijednosti drugih mreža. Za jednostavne perceptrone nema razlike između jednostrane komunikacije (učenje) i obostrane komunikacije (sinkronizacija).

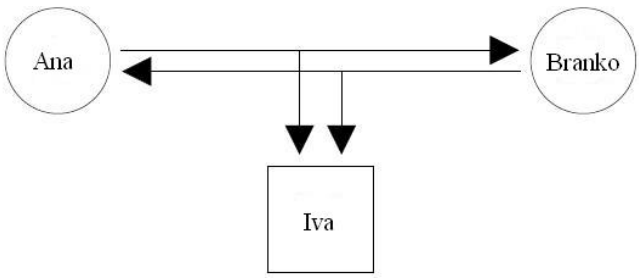
Sinkronizacija neuronskih mreža je poseban slučaj *online* algoritma. Dvije neuronske mreže kreću sa slučajno odabranim vektorima težina. U svakom koraku, mreže primaju zajednički ulazni vektor, računaju izlaznu vrijednost te ju šalju jedna drugoj. Ako se slažu u preslikavanju vrijednosti trenutačnog ulaza na izlaz, težine se mijenjaju u skladu sa prikladnim pravilom učenja. U slučaju diskretnih vrijednosti težina ovaj proces vodi do potpune sinkronizacije u konačnom broju koraka. [3][11]

III. PROTOKOL ZA NEURONSKU RAZMJENU KLJUČA

Pri analizi protokola za neuronsku razmjenu ključa vrijede neke uobičajene pretpostavke kod analize kriptografskih sustava:

- Napadač Iva zna sve poruke razmijenjene između Ane i Branka, tako svatko jednaku količinu informacija o svim drugima. Također, sigurnost komunikacije ne ovisi o nekom posebnom svojstvu informacijskog kanala.
- Iva nije u mogućnosti mijenjati poruke, tj. može izvoditi samo pasivne napade.
- Algoritam je javan.

Na slici 2. je prikazan koncept razmjene ključa između Ane i Branka, te pasivni napadač Iva koja prisluškuje javni kanal [3][6].

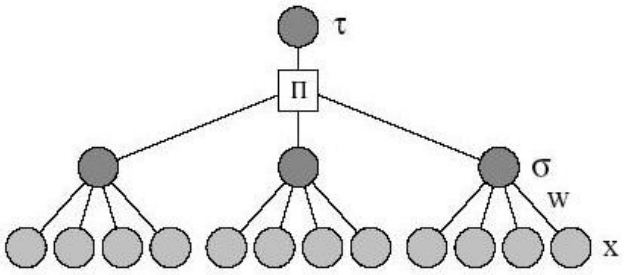


Slika 2. Prikaz protokola za razmjenu ključa

B. TPM mreže

TPM (engl. *Tree Parity Machine, TPM*) mreže spadaju u višeslojne neuronske mreže bez povratne veze. Takav tip neuronske mreže koriste sve strane u osnovnim protokolima temeljenim na neuronskoj kriptografiji. TPM mreže imaju svojstvo da je sinkronizacija obostranim učenjem puno brža od jednostranog učenja.

Na slici 3. je prikazana TPM neuronska mreža. Mreža se sastoji od *K* skrivenih neurona – perceptrona sa nezavisnim poljima. Svaki perceptron sadrži *N* ulaznih neurona te jedan izlazni neuron.



Slika 3. Prikaz TPM mreže

Sve ulazne vrijednosti su binarni brojevi

$$x_{i,j} \in \{-1, 1\} \tag{2}$$

Težine, koje određuju preslikavanje ulaza na izlaz su diskretne vrijednosti u rasponu od *–L* do *L*.

$$w_{i,j} \in \{-L, \mathrm{L}\} \tag{3}$$

Index *i* označava *i*-ti skriveni neuron u TPM mreži, a *j* označava element vektora.

Kao i u drugim neuronskim mrežama, suma trenutačnih ulaznih vrijednosti se koristi za izračun izlazne vrijednosti skrivenih neurona. Svaki skriveni neuron je potpuno određen svojim lokalnim poljem:

$$h_i = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} w_{i,j} x_{i,j} \tag{4}$$

Izlaz σ_i i-tog skrivenog neurona je definiran kao

$$\sigma_i = \mathrm{sgn}\,(h_i) \tag{5}$$

Kako bi se osigurala binarna izlazna vrijednost, ako h_i poprimi vrijednost *0* onda se ta vrijednost preslikava u σ_i *=-1*.

Ukupna izlazna vrijednost τ je određene umnoškom skrivenih neurona.

$$\tau = \prod_{i=1}^{K} \sigma_i \tag{6}$$

Iz izraza (6) se može vidjeti da τ pokazuje je li ukupan broj skrivenih neurona sa vrijednošću σ_i *=-1* paran ili neparan.

Na početku sinkronizacijskog procesa Anina i Brankova TPM mreža kreću sa slučajno odabranim i time nepovezanim vektorima težina. U svakom koraku sinkronizacije *K* ulazni vektori x_i su slučajno generirani te se računa odgovarajuća izlazna vrijednost $\tau^{A/B}$.

Ana i Branko šalju izlazne vrijednosti svojih mreža jedan drugome, ako se izlazne vrijednosti ne slažu $\tau^A \neq \tau^B$ tada se ni težine ne mijenjaju. Inače, težine se mijenjaju prema jednom od pravila učenja prikladnog za sinkronizaciju:

- Hebbianovo pravilo učenja – obje neuronske mreže uče jedna od druge.

$$w_{i,j}^{+} = g(w_{i,j} + x_{i,j}\tau\,\theta(\sigma_i\tau)\theta(\tau^A\tau^B) \tag{7}$$

- Anti-Hebbianovo pravilo učenja – mreže se uče na temelju vrijednosti suprotnih od njihovih izlaznih vrijednosti.

$$w_{i,j}^{+} = g(w_{i,j} - x_{i,j}\tau\,\theta(\sigma_i\tau)\theta(\tau^A\tau^B) \tag{8}$$

- Pravilo učenja slučajne šetnje – izlazna vrijednost nije važna dok je god ista za sve neuronske mreže koje sudjeluju.

$$w_{i,j}^{+} = g(w_{i,j} + x_{i,j}\,\theta(\sigma_i\tau)\theta(\tau^A\tau^B) \tag{9}$$

Primjenom bilo kojeg od tih pravila učenja mijenjaju se samo težine, koje su povezane sa skrivenim neuronima gdje je $\sigma_i = \tau$. Time je nemoguće doći do informacije koje su težine promijenjene ako se ne zna točna konfiguracija neuronske mreže.

Pravila učenja moraju osigurati da težine ostanu unutar granica *–L* i *L*. To se postiže funkcijom *g(w)* unutar svakog pravila učenja.

$$g(w) = \begin{cases} \mathrm{sgn}(w)L \;\; gdje\;\mathrm{je}\;|w| > L \\ \\ \mathrm{w,\,inace} \end{cases} \tag{10}$$

U protokolu neuronske razmjene ključa se primjenjuje neuronska sinkronizacija. Ana i Branko moraju koristiti TPM mrežu sa jednakom unutrašnjom strukturom. Parametri *K*, *L*, i *N* su javno poznati. Svaka neuronska mreža kreće sa slučajno odabranim težinskim vektorima. Inicijalni težinski vektori su tajni. Tijekom procesa sinkronizacije samo ulazni vektori x_i i izlazne vrijednosti τ^A, i τ^Bsmiju biti prenošene javnim kanalom. Ana i Branko smiju znati samo strukturu svojih TPM mreža. Razlika između obostranog i jednostranog učenja je ključna za sigurnost ovog protokola. Kako Iva ne može utjecati na Aninu i Brankovu mrežu ona u većini slučajeva ne može postići sinkronizaciju prije nego što Ana i Branko završe stvaranje tajnog ključa te prestanu sa prijenosom podataka kroz javni kanal. [3][11][12]

C. Primjer rada protokola

Ana i Branko imaju svaki svoju TPM mrežu sa slučajno generiranim vektorima težina. Kreira se zajednički slučajno odabrani ulazni vektor vrijednosti. Svatko računa vrijednost skrivenih neurona svoje mreže te ukupnu izlaznu vrijednost. Ana i Branko javnim kanalom uspoređuju izlazne vrijednosti te ako su različite tada se vraćaju na korak stvaranja slučajno odabranih ulaznih vektora vrijednosti. Ako su izlazne vrijednosti jednake prelazi se na primjenu jednog od prikladnih pravila učenja.

Postoje tri tipa koraka u procesu:

- Atraktivni – imaju jednake vrijednosti skrivenih neurona.
- Repulzivni – imaju jednake izlazne vrijednosti ali različite vrijednosti skrivenih neurona.
- Pasivni- imaju različite izlazne vrijednosti.

Nakon što je postignuta puna sinkronizacija Ana i Branko mogu koristiti tako definirane vrijednosti vektora težina kao zajednički tajni ključ. Nakon što jednom postignu potpunu sinkronizaciju mreže ostaju sinkronizirane.

D. Protokol za razmjenu ključa koji uključuje autentifikaciju

U kriptografiji, autentifikacija predstavlja jednako važnu cjelinu kao i sigurna razmjena ključeva. Jednosmjerne funkcije sažimanja se mogu jednostavno upotrijebiti, ali one nemaju izravne poveznice sa konceptom neuronske kriptografije.

Sinkronizacija TPM mreža obostranim učenjem uspijeva samo ako su sve mreže dobile iste ulazne vektore. To načelo se može primijeniti u implementaciji mehanizma autentifikacije u protokolu za neuronsku razmjenu ključa. Da bi to postigli, Ana i Branko posjeduju odvojene ali identične generatore pseudoslučajnih brojeva. Kako se generatori inicijaliziraju sa sjemenom koju znaju samo Ana i Branko, dati će jednaki izlazni slijed bitova koji se onda koristi za kreiranje ulaznih vektora x_i. Time Ana i Branko mogu sinkronizirati svoje mreže bez razmjene ulaznih vrijednosti javnim kanalom. Ivi nije poznata vrijednost upotrijebljena pri inicijalizaciji generatora te ne može sinkronizirati svoju mrežu zbog nedostatka informacija o ulaznim vektorima.

Potreba za zajedničkim tajnim ulaznim vektorom ne predstavlja bitnu prepreku u protokolu jer i drugi autentifikacijski protokoli zahtijevaju postojanje zajedničkih osnovnih podataka.[6]

Posljedično, postizanje pune sinkroniziranosti dokazuje da i Ana i Branko znaju sjeme te se time može provesti autentifikacija. [3]

IV. POZNATI NAPADI NA PROTOKOL

Glavni problem napadača Ive je činjenica da joj unutrašnje strukture TPM mreža Ane i Branka nisu poznate. Kako promjena težina w ovisi o vrijednostima σ_i, da bi napad bio uspješan potrebno je točno pogoditi stanje skrivenih neurona.

Većina poznatih napada koristi taj pristup.

A. Napad grubom silom

Da bi takav napad uspio, napadač mora ispitati sve moguće vrijednosti ključa (sve moguće vrijednosti težina $w_{i,j}$). Ako se mreža sastoji od K skrivenih neurona, $K*N$ ulaznih neurona te su granice vrijednosti težina od $-L$ do L to daje $(2L+1)^{KN}$ mogućih kombinacija. Prikladnim odabirom tih parametara takav napad nije moguće izvesti. [12]

B. Jednostavni napad

U ovom tipu napada Iva trenira svoju TPM mrežu sa vrijednostima koje se sastoje od ulaznih vektora x_i i izlazne vrijednosti τ. Ti podaci su dostupni jednostavnim prisluškivanjem javnog kanala dok Ana i Branko komuniciraju. TPM mreža napadača ima jednaku strukturu kao i TPM mreže Ane i Branka te u prvom trenutku ima slučajno odabrane ulazne vektore, kao i mreže Ane i Branka. U svakom koraku Iva računa izlaznu vrijednost svoje TPM mreže. Dalje, Iva koristi ista pravila učenja kao i Ana i Branko ali je vrijednost τ^E zamijenjena vrijednošću τ^A. U ovom napadu Iva koristi svoju mrežu da dođe do saznanja o strukturi Anine mreže, čak i ako je ukupna izlazna vrijednost različita. [3]

C. Geometrijski napad

Ova vrsta napada pokazuje bolja svojstva nego jednostavan napad jer Iva uzima u obzir i vrijednosti τ^E i vrijednosti lokalnih polja skrivenih neurona svoje mreže. Kao i u jednostavnom napadu Iva pokušava imitirati Branka ali nema mogućnosti komunikacije sa Anom. Dok je $\tau^A = \tau^E$ to se može ostvariti jednostavnom primjenom istog pravila učenja koje koriste Ana i Branko. Kada je $\tau^E \neq \tau^A$ Iva ne može spriječiti Anu u promijeni vrijednosti težina svoje mreže. Umjesto toga Iva pokušava ispraviti stanja skrivenih neurona svoje mreže korištenjem vrijednosti lokalnih polja kao dodatnom informacijom. Vrijednosti lokalnih polja mogu ukazati na stupanj povjerenja povezan sa izlaznom vrijednošću skrivenih neurona. Niska apsolutna vrijednost ukazuje na visoku vjerojatnost da je $\sigma_i^A \neq \sigma_i^B$ te Iva mora promijeniti izlaznu vrijednost σ_i^E skrivenog neurona sa minimalnom vrijednošću lokalnog polja $|h^E_i|$ i ukupnu izlaznu vrijednost τ^E prije primjene pravila učenja. Ovaj napad predstavlja najuspješniji poznati napad ako napadač koristi samo jednu TPM mrežu. [3]

D. Napad većinskim brojem rješenja

U ovom napadu Iva koristi M TPM mreža. Na početku procesa sinkronizacije težinski vektori w svih Ivinih mreža su odabrani slučajno. Kao i u drugim napadima Iva ne mijenja vrijednosti težina u koracima gdje je $\tau^E \neq \tau^A$. Kada je $\tau^E = \tau^A$ Iva računa izlaznu vrijednost $\tau^{E,m}$ svoje TPM mreže, ako je ta vrijednost različita od vrijednosti τ^A, Iva traži skriveni neuron na poziciji i koji ima minimalnu apsolutnu vrijednost lokalnog polja. Tada se izlazne vrijednosti $\sigma_i^{E,m}$ i $\tau^{E,m}$ invertiraju slično kao i u geometrijskom napadu. Iva tada prebrojava $\sigma_i^{E,m}$

vrijednosti svojih TPM mreža i odabire ono koja se najčešće pojavljuje. Taj prikaz usvajaju sve Ivine neuronske mreže za upotrebu pravila učenja. [3]

E. Napad genetskim algoritmom

Ovaj napad ne ovisi o optimiranju predviđanja vrijednosti σ_i već je zasnovan na evolucijskom algoritmu. Iva uobičajeno kreće sa jednom TPM mrežom sa slučajno odabranim vrijednostima ali može koristiti do M TPM mreža.

Kada Ana i Branko mijenjaju vrijednosti težina svojih mreža zbog ispunjenja uvjeta $\tau^A = \tau^B$ primjenjuje se sljedeći genetski algoritam:

- Dok Iva ima najviše $M/2^{K-1}$ TPM mreža, ona određuje svih 2^{K-1} vrijednosti σ_i koje oponašaju izlaznu vrijednost τ^A. Kasnije se te vrijednosti koriste za promjenu težinskih vektora svih Ivinih mreža u skladu sa pravilom učenja. Time Iva stvara 2^{K-1} varijanti svake TPM mreže u ovom koraku (mutacijski korak)
- Ako Iva već ima više od $M/2^{K-1}$ neuronskih mreža, zadržava se samo najuspješnija TPM mreža. To se postiže odbacivanjem svih mreža koje su predvidjele manje od U izlaznih vrijednosti τ^A u zadnjih V koraka učenja, gdje je $\tau^A = \tau^B$. Ovaj korak predstavlja selekcijski korak.

Efikasnost napada genetskim algoritmom najviše ovisi o algoritmu koji bira najuspješniju neuronsku mrežu. [4]

Od gore spomenutih napada, napadi većinskim brojem rješenja, i napad genetskim algoritmom predstavljaju najefikasnija rješenja, gdje je napad genetskim algoritmom bolji jedino u slučaju da L nije prevelik.

Ana i Branko teoretski mogu postići bilo koji stupanj sigurnosti povećanjem sinaptičke dubine svojih mreža ali to uzrokuje i povećanje vremena potrebnog za sinkronizaciju mreža.

.Postoji, naravno i mogućnost razvoja neke nove metode napada koja će uništiti sigurnost neuronske kriptografije ali to se može reći i za svaki drugi sustav osim jednokratne bilježnice.

V. ZAKLJUČAK

Neuronska kriptografija predstavlja mladu i relativno nepoznatu granu kriptografije. Ta vrsta kriptografije pokazuje određeni potencijal koji se može uspješno primijeniti u raznim sustavima. Ali opet, barem za sada, nema niti jedne praktične primjene neuronske kriptografije.

Zašto je to tako? Može se reći da postoji određeni kulturalni problem u kriptografiji – dok god postojeće metode dobro rade, tj. nema uspješnih napada, nitko ne želi koristiti nove metode. Naravno, postoji i mogućnost da neuronska kriptografija pokaže razočaravajuće rezultate u praktičnoj primjeni. Za neke iscrpnije pokazatelje nužno je upotrijebiti protokol u stvarnim aplikacijama. Jedno područje koje se čini logičnim za implementaciju neuronske kriptografije su ugrađeni sustavi koji posjeduju ograničenu računalnu snagu. Zahvaljujući činjenici da neuronska sinkronizacija zahtijeva samo osnovne matematičke operacije radi se i na sklopovskoj potpori neuronskim mrežama za potrebe kriptografije.

LITERATURA

[1] A. Dujella i M. Maretić, Kriptografija, Element, Zagreb, 2007., (na hrvatskom)

[2] A. Klimov, A. Mityaguine, and A. Shamir, Analysis of Neural Cryptography, Advances in Cryptology - ASIACRYPT, edited by Y. Zheng (Springer, Berlin, 2003), pp. 288, 2002.

[3] A. Ruttor, Neural Synchronization and Cryptography , PhD thesis, 2006.

[4] A. Ruttor, W. Kinzel, R. Naeh, and I. Kanter, Genetic attack on neural cryptography, Phys. Rev. E 73, 036121, 2006.

[5] B. Kröse and P. van der Smagt, An Introduction to Neural Networks, University of Amsterdam, 8th edition, 1996.

[6] B. Schneier, Applied Cryptography, John Wiley & Sons, 2nd edition, 1996.

[7] E. Klein, R. Mislovaty, I. Kanter, A. Ruttor, and W. Kinzel, Synchronization of neural networks by mutual learning and its application to cryptography, Advances in Neural Information Processing Systems 17, MIT Press, Cambridge, 2005.

[8] P. Peretto, An Introduction to the Modeling of Neural Networks, Cambridge University Press, Great Britain, 1992.

[9] S. Lončarić, predavanja iz predmeta Neuronske mreže, Fakultet elektrotehnike i računarstva, 2008., (na hrvatskom)

[10] S. Picek i M. Golub: Kvantna kriptografija: razvoj i protokoli. Proceedings Vol. V., DE & ISS & miproBIS & LG & SP, MIPRO 2009, Opatija, Hrvatska, pp. 122-127, 2009., (na hrvatskom)

[11] W. Kinzel, and I. Kanter, Neural Cryptography, 9th International Conference on Neural Information Processing, Singapore, 2002.

[12] Internet stranice:
http://en.wikipedia.org/wiki/Neural_cryptography (10.02.2010.)
http://en.wikipedia.org/wiki/Perceptron (10.02.2010.)
http://www.learnartificialneuralnetworks.com/ (10.02.2010.)