About the author

Emil Eifrem is CEO of Neo Technology and co-founder of the Neo4j project (http://neo4j.com). Before founding Neo, he was the CTO of Windh AB, where he headed the development of highly complex information architectures for enterprise content management systems. Committed to sustainable open source, Eifrem sees his role at Neo as steering a balanced path between free availability of powerful graph database solutions and enterprise-level options where mission-critical capability is required. He is a frequent conference speaker and a well-known author and blogger on NoSQL and graph databases.

References

1. 'The Financial Cost of Fraud'. PKF. Accessed Mar 2016. www.pkf-littlejohn.com/the-financial-cost-of-fraud-2015.php.
2. 'Bust-out fraud'. Experian, 2009. Accessed Mar 2016. www.experian.com/assets/decision-analytics/white-papers/bust-out-fraud-white-paper.pdf.
3. Ryle, Gerard et al. 'Banking giant HSBC sheltered murky cash linked to dictators and arms dealers'. ICIJ, 8 Feb 2015. Accessed Mar 2016. www.icij.org/project/swiss-leaks/banking-giant-hsbc-sheltered-murky-cash-linked-dictators-and-arms-dealers.

Online recruitment services: another playground for fraudsters

Sokratis Vidros

Constantinos Kolias

Georgios Kambourakis

Sokratis Vidros, University of the Aegean, Greece; Constantinos Kolias, George Mason University, US; Georgios Kambourakis, University of the Aegean

It is increasingly common for someone actively seeking a job online to come across appealing but fake ads offering high wages, flexible hours, teleworking and career growth opportunities. Usually, job ads with such favourable conditions are examples of Online Recruitment Fraud (ORF) which attempt to collect unsuspecting candidates' personal information. ORF is a relatively new field of variable severity that can escalate quickly to extensive scams.

A survey conducted by FlexJobs in 2015, revealed that for every legitimate job posting, there were around 60 fraudulent ones, yet only 48% of applicants stated they were even aware of employment scams while searching for new career opportunities online.[1] In addition, 7% of job seekers have been victims of employment scams at least once, despite warnings from the FBI and Better Business Bureau. Workable, a widely-used online recruiting software, reported that well-crafted fraudulent job ads for blue collar or secretarial positions in densely populated countries can collect up to 1,000 resumés per day.[2]

Interestingly, in 2012 a job seeker received more than 600 resumés in one day after posting a fake job ad on Craigslist in order to identify his competitors.[3] That same year, the Australian Bureau of Statistics published a report about personal fraud stating that 6 million people were exposed to several forms of scam, including employment scams, during any given year.[4] The report concluded that the Australian economy has been severely affected by this exposure.

Moving to the cloud

Corporate hiring has recently moved to the cloud. LinkedIn was the first service to demonstrate the huge opportunity that lay in building online recruiting tools. Dozens of companies followed and built online automated systems – Applicant Tracking Systems (ATS) – to help organisations recruit talent and job seekers to find jobs. These tools made the hiring process more immediate, accurate and cost efficient.

"Online hiring services attracted the interest of scammers, whose malicious behaviour aiming to steal personal information both inflicts economic damage and harms the reputation of the ATS stakeholders"

On the downside, the increasing adoption of ATS has also led to more incidents

of ORF. This is because online hiring services attracted the interest of scammers, whose malicious behaviour aiming to steal personal information both inflicts economic damage and harms the reputation of the ATS stakeholders. Scammers creatively and persistently try to manipulate the functions of these systems. Most frequently they do so by capitalising on the job syndication process, which is also referred to as employment scamming or job scamming.

More specifically, ATS enforces a streamlined workflow that starts with the composition of the job description and continues with the dissemination of the job ad to popular job boards (eg, Indeed, Monster, CareerBuilder, Simply Hired, etc) as well as social networks (eg, LinkedIn, Facebook, Twitter, etc).[5] For every incoming resumé, the utilised ATS attempts to construct a complete candidate profile including detailed information about the candidate's education, work experience, skills, social connections and professional network. Additional information is also collected from multiple sources such as the Application Programming Interfaces (API) of popular social networks, web crawlers and automated scrapers. The typical lifecycle of ATS is shown in Figure 1.

The most popular approaches used by scam practitioners can be categorised into two main groups. The first pertains to fake job advertisements that aim at collecting sensitive information and luring the candidate into filling application forms which usually request the full name, phone number, address and postal code of the candidate. As a result, they are able to build entire databases of private information provided to them by users who unknowingly applied to non-existing positions.

In fact, more sophisticated scammers go as far as obtaining the educational and working experience of their victims as well as their online social footprint. These comprehensive databases can then be sold to third parties such as cold-callers, marketeers, proponents of political campaigns and sometimes even website administrators planning on sending out targeted bulk emails including links to generate much-wanted page views. The exploitation of these databases by more advanced evil-doers, including botnet operators planning to spam the users, is also a possibility.

Figure 1: The ATS lifecycle.

Identity theft

The second approach, however, is far more sinister as it usually escalates to identity theft associated with economic trickery. In these cases, scammers may pretend to be either legitimate or fictional employers and use the ATS as a means to advertise fake jobs. In their effort to maintain a veneer of legitimacy and respectability, they go as far as organising fake aptitude tests and scheduling interviews to ensure they gain the candidates' trust and engage with them.

Once the trust has been established, they convince their victims to provide highly sensitive information ranging from their Social Security numbers, ID numbers and a copy of their passport which, needless to say, can be used for endless types of fraud. Some social engineering-powered scammers may even try to trick their candidates into directly filling out fake deposit forms in order to steal their bank account details or straight-out ask them to wire money under the guise it will be put towards necessary visa or work travel expenses.

"Some social engineering-powered scammers may even try to trick their candidates into directly filling out fake deposit forms in order to steal their bank account details"

A possible third case related to ORF is the endorsement of fake jobs by legitimate companies as a means to survive internal audits. This may sound questionable, but companies may very well opt to spend extra funds on the virtual recruitment of positions they do not need to fill in order to justify expenses. Though it may be frustrating for the unassuming victims who apply to practically non-existent jobs, there is also a side benefit to this dubious practice and that is that the company acquires a talent pool of candidates that they might consider for future vacancies. Such virtual job openings are practically undetectable through the use of automated tools as they are valid posts made by a legitimate company.

Lastly, the act of automated replies to legitimate posts by fake applicants may be less common but is also regarded as a form of online recruitment abuse. Ill-motivated individuals try to suffocate the databases of the 'victim' company, thus making it harder to rely on their online systems or may even cause Denial of Service (DoS) by exhausting the ATS resources. Additionally, mass applications can be mitigated by well-known bot detection mechanisms such as Captcha.

The repercussions of these last two cases do not pose a serious threat to the stakeholders of ATS and are considered outside the scope of this article.

Timeliness and severity

It is a fact that ATS solutions are increasingly becoming an essential tool used in hiring. Recent statistics indicate that nearly 75% of the Fortune 500 companies are using some type of ATS to both review and manage all their incoming applicants.[6] Meanwhile, small and medium-sized businesses (SMBs) are starting to purchase modern ATS to streamline their hiring process. Added to that, a significant increase in ATS usage is expected in the near future, since around 50% of potential employers currently use manual methods to handle their hiring.[7]

"Job scammers intend to take advantage of ATS solutions' syndication features and use them as the primary channel for distributing phony content in high volumes. This way they are able to lure a large number of victims"

An ATS provides an efficient, automated and systematic way to manage the growing number of applications. According to Bloomberg, there were approximately 5 million job vacancies in the US in 2015, most of which were advertised online.[8] Moreover, 71% of the workforce are either actively pursuing or are at least open to new job opportunities using web resources.[9] This figure is only reinforced by the fact that every year around 40 million people change jobs in the US and around 30% of all Google search requests, in other words about 300 million per month, are employment related.[10] Without ATS, the amount of time and paperwork required to review each candidate profile would be immense, especially for large organisations that may receive up to 1,000 resumés per week.

In the meantime, social recruiting is also on the rise. Most recruiters already use or plan to use social media for scouting talent. Jobvite's social recruiting survey for 2013 indicated that 78% of the industry reviewed a candidate's social profile before making a hiring decision.[11] In fact, it is worth pointing out that employers who capitalised on social media to recruit found a 49% improvement in candidate quality. With 92% of surveyed recruiters reporting that they have hired someone through LinkedIn, the latter appears to be in the lead as the most popular social media used for recruitment, with Facebook and Twitter coming in second and third place.[12,13]

Clearly, ATS and social media play a key role in modern recruiting. On the one hand, the interoperability of such platforms enables organisations to hire employees effectively, but on the other hand, it introduces multiple vulnerabilities that could be exploited and invoke issues in terms of data privacy, data leakage and identity theft.

Typically, job scammers intend to take advantage of ATS solutions' syndication features and use them as the primary channel for distributing phony content in high volumes. This way they are able to lure a large number of victims. In most cases, they target candidate profiles of diverse background so as to create more complete and thus expensive databases. The risk of an employment scam is higher for younger employees as they have a larger presence in social networks compared to older ones. Moreover, telecommuting is becoming more popular for white collar professionals (according to Gallup, nowadays nearly 37% work remotely, four times as many as in 1995) and young people often prefer to work from home.[14]

Essentially, the majority of fraudulent job ads found online concerns overpaid work-from-home positions. Career advisors, hiring managers and content writers in recruiting firms try to warn job seekers on ORF practises by publishing articles describing guidelines candidates should follow in order to spot a fake posting (ie, appealing, home-sourced offers should be scrutinised). Furthermore, a plethora of job boards provide reporting tools where candidates can expose scam practitioners. But apart from these informal sources, there is a distinct lack of relevant scientific literature on employment scams.

Similarities and differences

So far, employment scam detection remains largely unexplored. At a high level, the problem can be defined as the process of classifying a raw corpus of job ads into fraudulent and legitimate. Such a process typically entails modelling information about the textual, structural and contextual characteristics of that content. By applying techniques such as Natural Language Processing (NLP) and machine learning classification, security researchers try to effectively evaluate the scam index of each entry. Generally, scam detection utilises supervised and semi-supervised classifiers based on regression trees, Support Vector Machines (SVM), Random Forests (RF), and neural networks.

At this point, the reader can easily notice that employment scam detection is similar to well-studied but still developing phenomena such as email spam, phishing, Wikipedia vandalism, cyber-bullying and trolling. Indeed, employment scam detection bears a resemblance to these problems as it mainly relies on the existence of textual elements – that is, words or phrases that convey the presence of an anomaly.

Conversely, the aforementioned relevant problems possess additional characteristics

that significantly aid the detection process and enable researchers to build clear-cut countermeasures that work on the transport and network layer. Such characteristics include dedicated communication protocols (ie, TCP/IP, HTTP, SMTP), social information and generated metadata. For example, email spam detection also relies on the abuse of the SMTP protocol (eg, address spoofing, invalid MIME header fields) or senders who are associated with the dissemination of large volumes of email.

Additionally, spam detection is benefited by the social characteristics of recipients such as their circle of contacts – that is, contacts of a user are less likely to send the user an unwanted message even if that message bears advertising content. Phishing content can be spotted via technical characteristics attesting unauthorised redirects to other domains, the level of visual or structural similarity among online services as well as previously reported bad user experience. Lastly, metadata deriving from user's online social behaviour and history, as well as timestamped revisions of the original data as used in Wikipedia vandalism detection, may prove crucial for detecting irregularities.

Impediments

Over the past few years, attackers have become rather resourceful. Phony job ads are becoming increasingly hard to distinguish from those that are legitimate, so countermeasures are usually ad hoc and their practical value is questionable. In the case of employment scams, non-textual information has a limited contribution to make to malicious content disclosure. Structural anomalies – including invalid HTML mark-up – are in most cases the products of low-skilled practitioners and serious attackers can easily circumvent them. Moreover, contextual attributes and metadata such as additional information about the location of a job or uploading the corporate logo are often neglected even by actual users.

At the same time, ATS solutions are web applications that typically work over the HTTP/HTTPS protocol. In other words, they do not entail any dedicated communication protocol that could possibly provide additional indications of an employment scam. Added to that, the malicious user may generate and publish a single job ad and then make no further interactions with the system. In cases when the assailant impersonates an existing (or a fictional) business rather than a single user it is even harder to deduce whether the advertised job postings are legitimate or not as any related online social data cannot lead to a succinct conclusion about their identity.

Primarily, scammers exploit the ability of ATS to broadcast newly created job ads to multiple sources in order to reach out to a large audience and collect many resumés fast. The automated job distribution is achieved through four channels:

- Daily RSS, XML or JSON feeds generated by the ATS and parsed by job posting sites.
- Direct API calls from the ATS to job boards.
- Sharable URLs that are redirected to the corresponding application form powered by the ATS.
- Dedicated email drop box addresses handled by the ATS bound to each position where any inbound emails containing resumés are handled as a new application.

Given all these, it is evident that the problem of employment scam detection is a non-trivial, primarily text-based problem. Thus, as already pointed out, existing solutions proposed for the aforementioned relevant topics cannot be easily adapted to the quirks of the problem. According to ATS vendors, the high similarity between the fraudulent and the legitimate content is the biggest obstacle when designing defensive mechanisms against employment scams. Needless to say that for job scammers the malicious content inherently aims to be as indistinguishable as possible from the legitimate one.

Of course, affected ATS solutions already build defences against job scams, but those countermeasures are not always adequate and they mostly depend on their sales policy. For example, online services that allow a free registration process generally try to employ in-house remedies such as requiring users to provide a corporate

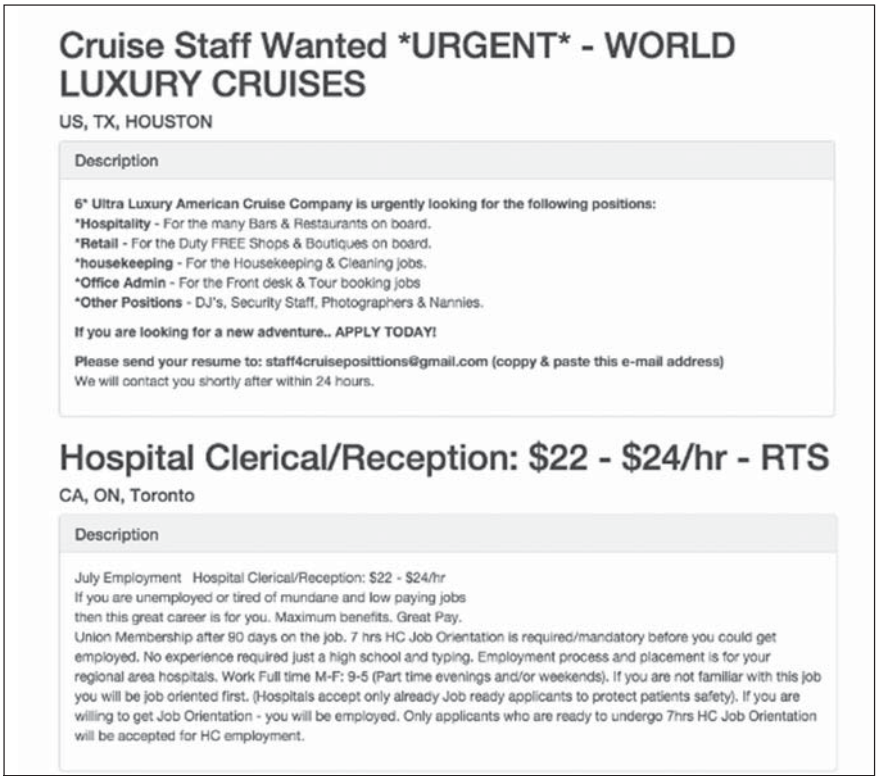

Figure 2: Examples of fraudulent job ads.

Feature type	Description
Text	Short job description.
	Poor company-related information.
	Frequent typos and grammar errors.
	Missing job requirements and benefits.
	Requirements for low-level degrees.
	Money symbols and figures, especially in job title.
	High ratio of capitalised characters.
	High ratio of consecutive punctuation.
	Suspicious words with spammy content.
	Active URLs redirecting to doubtful websites.
	Multiple email addresses inside the job ad.
	Prompts for external application process outside the ATS.
HTML	High emphasised word ratio (eg, b, em and strong HTML tags).
Metadata	Missing corporate logo.
	Missing screening questions.
	Telecommuting.

Table 1: Empirical rules denoting spammy content.

email address upon signup in order to verify their address through DNS Mail Exchanger (DNS MX) lookups to prevent spoofing attempts. On the other hand, an ATS with a restricted registration process has no particular need for fraud detection systems since every potential customer must go through a sales representative who will verify their profile and business.

Experimentation on real-life data

In our effort to investigate further employment scams and acquire practical knowledge on the peculiarities of the problem, we examined an exploratory corpus of 17,000 real-life job ads that were processed by Workable, a widely-used, cloud-based recruiting software. Dataset entries were manually reviewed and annotated as legitimate or fraudulent by expert users – that is, 95% of the entries were labelled as legitimate and 5% were marked as fraudulent (Figure 2). The decisions for the annotation process were based on Workable's internal processes, customer complaints and blatant ATS abuse.

A job ad encapsulates the details of the job opening and prompts candidates to apply. In short, it contains: the job title; the job description; the detailed list of benefits; skills and requirements candidates should meet; the geographical location of the opening; the job metadata (ie, the required experience level, the related industry and the salary range); and the company-related information such as the corporate website. Legitimate job ads are advertised at popular job boards, including Indeed, CareerBuilder, LinkedIn, Simply Hired and social networks, including Twitter, Facebook and Google+.

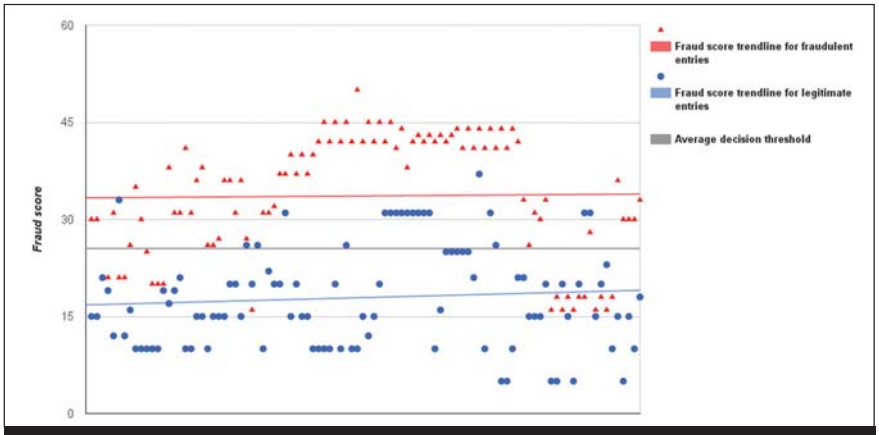

Figure 3: Fraud score distribution based on the statistical analysis of the preliminary hypotheses.

After examining the corpus, we established a preliminary list of empirical rules denoting spammy content (Table 1). As a next step, we statistically analysed the data set and we evaluated the predictive power of our hypotheses by defining a decision threshold computed on the total fraud score of each entry.

In further detail, each rule was attributed a scoring factor according to its entropy value computed on the dataset. The total fraud score for each entry, was calculated as the sum of the score of every rule the entry satisfies. Figure 3 summarises the process for a balanced sample of 200 entries and displays the estimated decision threshold for that sample. The results solidify the initial assumptions and encourage the design of a fully-fledged job ad classifier trained on similar features.

Conclusion

Employment scam detection is a difficult yet a very relevant and timely problem that demands further research efforts. It may still be in its infancy, but given the increasing volume of scam content and the extensive repercussions it can lead to, it is an intriguing topic that needs to be better understood and formulated. The absence of relevant tools and datasets may act as a roadblock to its fast resolution, thus the process of collecting and analysing recent job ad data will arm the research community with meaningful insight to the problem.

About the authors

Sokratis Vidros is a senior software engineer at Workable (www.workable.com) and a postgraduate research fellow at Info-Sec-Lab, University of the Aegean, Samos, Greece. His research interests include web application security and fraud detection. He is a graduate of the Electrical and Computer Engineering

school of the National Technical University of Athens, Greece. In 2013 he received his Master of Science in Communication and Computer Security from Telecom ParisTech and Institute Eurecom, France. Contact him at sokratis@workable.com.

Constantinos Kolias is a research assistant professor at George Mason University, Virginia, US. His research interests include security for 4G/5G communication protocols, wireless intrusion detection and IoT/M2M security. He received a PhD in computer science from the University of the Aegean. Contact him at kkolias@gmu.edu.

Georgios Kambourakis is an associate professor at the Department of Information and Communication Systems Engineering, University of the Aegean and the director of Info-Sec-Lab. His main research interest is in the field of mobile and wireless networks security and privacy. He has published over 100 articles in leading journals and conferences and he has been involved in several national- and EU-funded R&D projects in the areas of information and communication systems security. More info at: www.icsd.aegean.gr/gkamb.

References

1. Howington, Jessica. 'Survey: More millennials than seniors victims of job scams'. Flexjobs, 12 Sep 2015. Accessed Nov 2015. www.flexjobs.com/blog/post/survey-results-millennials-seniors-victims-job-scams.
2. 'Workable, Online recruiting software', Workable. Accessed 12 Dec 2015. www.workable.com.
3. Auld, Eric. 'Man posts fake job on craigslist, gets 600+ resumés'. Chemjobber, 2 Aug 2012. Accessed Sep 2015. http://chemjobber.blogspot.gr/2012/08/man-posts-fake-job-on-craigslist-gets.html.
4. 'Personal fraud, 2010-2011', Australian Bureau of Statistics. Accessed Sep 2015. www.abs.gov.au/AUSSTATS/abs@.nsf/mf/4528.0.
5. Del Castillo, Christine. 'All the best places to post your jobs'. Workable Blog, 4 May 2015. Accessed Nov 2015. http://blog.workable.com/best-places-post-jobs.
6. Peggs, Michael. 'Applicant tracking systems solved'. Peggs, 14 Jan 2015. Accessed Nov 2015. www.michaelpeggs.com/applicant-tracking-systems-solved.
7. Westfall, Brian. 'Human resources software buyer report'. Software Advice, 2015. Accessed Nov 2015. www.softwareadvice.com/resources/hr-buyer-report-2015/.
8. Stilwell, Victoria. 'There are now more than five million job openings in America'. Bloomberg, 10 Feb 2015. Accessed Nov 2015. www.bloomberg.com/news/articles/2015-02-10/job-openings-in-u-s-rose-by-181-000-in-december-to-5-03-million.
9. 'Jobvite job seeker nation study'. Jobvite, 2014. Accessed Sep 2015. http://web.jobvite.com/rs/jobvite/images/2014%20Job%20Seeker%20Survey.pdf.
10. McClure, Jennifer. 'Hiring and onboarding done right'. 25 Jul 2012. Accessed May 2015. www.slideshare.net/jennifermcclure/hiring-onboarding-done-right-nky-chambernky-shrm-7-24-2012.
11. 'Social recruiting survey for 2013'. Jobvite, 2013. Accessed Nov 2015. http://web.jobvite.com/rs/jobvite/images/Jobvite_2013_SocialRecruitingSurveyResults.pdf.
12. 'Social recruiting survey for 2014'. Jobvite, 2014. Accessed Sep 2015. www.jobvite.com/wp-content/uploads/2014/10/Jobvite_SocialRecruiting_Survey2014.pdfweb.jobvite.com/rs/jobvite/images/2014%20Job%20Seeker%20Survey.pdf.
13. Akiode, Segun. 'The social recruiting pocket guide'. Socialmeep, 19 Jun 2013. Accessed Nov 2015. http://socialmeep.com/infographic-the-social-recruiting-pocket-guide.
14. Jones, Jeffrey M. 'In U.S., telecommuting for work climbs to 37%'. Gallup 19 Aug 2015. Accessed Nov 2015. www.gallup.com/poll/184649/telecommuting-work-climbs.aspx.

Highwaymen to hackers


Rolf von Roessing

Rolf von Roessing, Forfa AG

Where once the scourge of the highways were cloaked highwaymen armed with single-shot pistols, our motorways and freeways are now under threat from a much more sinister menace – the Internet car hacker, armed with laptop and code.

The latest development, a product recall of more than one million vehicles in the US following the recent high-profile hack of a Jeep Cherokee by Charlie Miller and Chris Valasek showed that, as vehicles encompass more and more digital, wireless and connected technologies, the concept of being able to hack them has graduated from a theoretical (although impractical), possibility, to a genuinely workable reality.

Rapid acceleration

For more than a century, the automobile has been an isolated machine of metal and motor, its single most important purpose the transportation of passengers from one location to another. But that is now shifting, and over the past decade the automotive industry has seen the most rapid acceleration of change since Henry Ford's famous assembly line invention revolutionised mass production back in 1913.