Template:  
- URL:   
- Title:   
- Authors:   
- Notes:

##Review

##Citation

- URL: http://www.ijiee.org/papers/280-N011.pdf

- Title: An Efficient Network Monitoring and Management System

- Authors: Rafiullah Khan, Sarmad Ullah Khan, Rifaqat Zaheer, and Muhammad Inayatullah Babar

- Notes: This describes using Nagios in detail, which could be INCREDIBLY useful later because it turns out Nagios has a JSON API. My chat bot + Nagios API could be an awesome way of getting incredibly detailed information to the admin. Worth exploring.

##Review

According to Khan et al. 2013, large organizations require fast and efficient network monitoring systems that reports to a network administrator via email or SMS as soon as a problem arises, with details of the problem and locations affected. They go on to explain the merits of Nagios, a network monitoring tool, and its role in their system. It is extremely important, they mention, that the system be essentially autonomous in operation, as in a large company manual monitoring is very difficult.

The paper contains some basic instruction for configuring a Nagios setup and defines several ways for the software to check the status of various servers and services, and could act as a good guide for somebody new to Nagios configuration. However, the paper does not explore additional ways of informing administrators of issues, nor does it compare other software that may have similar features.

The method used to set up Nagios means that their system interacts with a Request Tracker, and as Nagios detects faults in the network it will send affected nodes plus other information via the Request Tracker as a ticket to the network admin. If the ticket isn’t resolved in an hour, the ticket is send to the second responsible network person. This method means that all persons are informed one by one until the ticket is marked as resolved. This could mean some dissonance may exist at times between different responsible persons, but this is not addressed in the paper.

###Citation

Rafiullah Khan, Sarmad Ullah Khan, Rifaqat Zaheer, and M. I. B. (2013). An Efficient Network Monitoring and Management System. *International Journal of Information and Electronics Engineering*, *3*(1), 122. https://doi.org/10.7763/IJIEE.2013.V3.280

- Url: http://www.jmlr.org/papers/volume12/collobert11a/collobert11a.pdf

- Title: Natural Language Processing (Almost) from Scratch

- Authors: Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukeuoglu, Pavel Kuksa

- Notes: Jason Weston was a research scientist at NEC Labs, Google and Facebook. That's some freaking pedigree right there.

##Review

No consensus has emerged whether a piece of software will ever be able to convert English text into a programmer friendly data structure that describes the meaning of the text, according to a paper by Collobert et al. 2011. The paper was written to accompany an attempt to build a natural language parser using a huge database of training data, and documenting the process of machine learning. Their approach was benchmarked using four standard NLP tests:

* Part of Speech tagging – This aims at labelling each word with a tag that indicates its *syntactic role*, such as plural, noun, adverb etc.
* Chunking (or shallow parsing) – This aims at labelling segments of a sentence with syntactic constituents such as noun or verb phrases, where each word is assigned a tag and encoded as a ‘begin-chunk’ or ‘inside-chunk’ tag.
* Named Entity Recognition – This aims to label elements into categories such as “PERSON” or “LOCATION”.
* Semantic Role Labelling – This aims at giving a semantic role to a syntactic constituent of a sentence.

Their experiments followed the standard evaluation procedure of the CoNLL challenges, a set of tasks with the goal of challenging the computer science community to create machine learning strategies which address proposed natural language processing problems.

The paper criticizes itself, noting that they used multilayer neutral networks, a 20-year-old technology, rather than something more modern, though they also note that the training algorithm used was only possible due to the tremendous progress in computer hardware. Due to their unique approach of trying to build from scratch rather than using work already established, much potentially relevant information from other papers and previous experiments could be construed as missing.

###Citation

Collobert, R., Weston, J., & Bottou, L. (2011). Natural language processing (almost) from scratch. *The Journal of Machine …*, *12*, 2493–2537. <https://doi.org/10.1145/2347736.2347755>

- URL: <http://researchbriefings.parliament.uk/ResearchBriefing/Summary/POST-PN-389>   
- Title: Cyber Security in the UK  
- Authors: Chandrika Nath  
- Notes: Potentially not very relevant, does it even count as a paper?

##Review

The Cyber Security in the UK paper explains, in detail, the British governments approaches to cyber security. It also describes various different types of attacks, such as data theft, attacks on critical information infrastructure, and attacks on physical infrastructure. The paper does well to inform the reader of terminology related to the field, such as the concept of air-gapping – network isolation, and zero-day attacks – previously unknown and unprotected attacks.

The paper suggests that common cyber security measures include methods such as the deployment of firewalls, using up to date anti-virus software, regular software patching, access management, encryption, and use of intrusion detection software. It also stresses the importance of security in industrial control systems, such as smart metering of gas and electricity consumption in homes, to avoid data falsification or damage to systems.

While there are some examples of major high profile attacks such as the Stuxnet virus and the data thefts at Lockheed Martin, the details of these attacks are very lacking and there are few sources to follow for more information. The paper offers no real scrutiny or analysis, and merely informs the reader, where it would have been nice to see a comparison between other countries cyber security plans.

##Citation

Chandrika, N. (2011). Cyber Security in the UK. *POSTnote*, (389), 1–4. Retrieved from http://www.parliament.uk/business/publications/research/briefing-papers/POST-PN-389

- URL: <http://ieeexplore.ieee.org/document/5375542/>   
- Title: Visualising Cyber Security: Usable Workspaces  
- Authors: Glenn A. Fink ; Christopher L. North ; Alex Endert ; Stuart Rose  
- Notes: Apparently command line tools are primitive, so we should all use GUI’s instead.

##Review

This paper provides an insight into an experiment involving adding more visualisations into cyber security, and phasing out ‘primitive’ command line tools. They replaced eight cyber analysts screen setups with a 4x2 configuration and recorded results when faced with generic visualisations of Net-flow and Snort alert data, which were met with mixed opinions.

The paper quotes the opinions of many analysts, even stating broadly that cyber analysts in general dislike visualisations, and prefer command line due to its flexibility and expressive power. While they did find several situations where visualisations helped in finding complex patterns in real world data, they found it very hard to sell the idea of visualisation to seasoned analysts. This distrust may stem from poor performance of instruction detection systems, which attempt to automate and “simplify” the process, as the number of false positives emitted is supposedly very high.

Much of the experimentation done in the paper was met with harsh response by the analysts, with many comments such as their original approaches, using grep or SQL queries, being considerably faster than the visualised equivalent. These comments are met with unconvincing defences by the paper, stating that is was not a fault of the visualisation tool in use, but rather bad database management. The paper could have spent much more time discussing possible alternatives rather than trying to defend its choices when met with criticism.

##Citation

Fink, G. A., North, C. L., Endert, A., & Rose, S. (2009). Visualizing cyber security: Usable workspaces. In *6th International Workshop on Visualization for Cyber Security 2009, VizSec 2009 - Proceedings* (pp. 45–56). https://doi.org/10.1109/VIZSEC.2009.5375542

- URL: <http://www.ijser.org/researchpaper/Study-of-Latest-Emerging-Trends-on-Cyber-Security-and-its-challenges-to-Society.pdf>   
- Title: Study of Latest Emerging Trends on Cyber Security and its challenges to Society  
- Authors: Ravi Sharma  
- Notes: Focuses on security emerging trends while adopting new technologies. Bachelor of engineering with a Hotmail email address, possibly an idiot.

##Review

This study shows new trends in cyber security based on the adoption of new technologies, claiming that due to the Windows 8 and onwards unified architecture between devices, attacks will be easier than ever between a range of systems. It also claims that due to this, somehow it would be possible to develop malicious applications like those for Android.

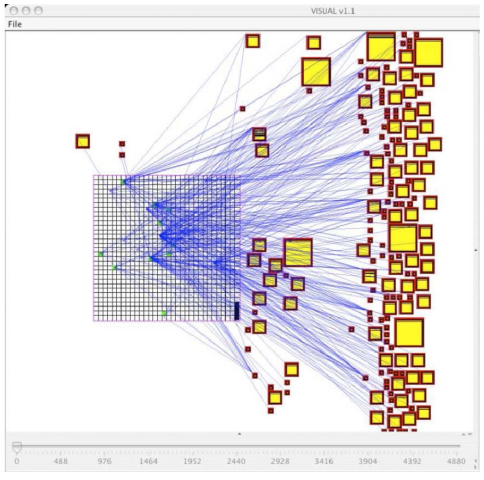
The study makes numerous claims but a good number seem to be not fully backed up by their referenced media, or lacks any sort of study that can be referenced at all. The study also claims via its abstract to discuss lack of coordination between security agencies and critical IT infrastructure, though this was not covered in detail.

##Citation

Sharma, R. (2012). Study of Latest Emerging Trends on Cyber Security and its challenges to Society. *International Journal of Scientific & Engineering Research*, *3*(6). Retrieved from http://www.ijser.org

- URL: <http://ieeexplore.ieee.org/document/6007132/>   
- Title: A Survey of Visualisation Systems for Network Security  
- Authors: H. Shiravi, A. Shiravi, A.Ghorbani  
- Notes: Also includes me fiddling with the new in-line citation style. Has a super-useful table of where information can come from, could be good to link to later when looking at security events from IDS and how to format data.

##Review

Security Visualization is a very young term, and many common visualisation techniques are not designed for security related data (Shiravi, Shiravi, & Ghorbani, 2012). Manually traversing textual logs is not only frustrating and time consuming, but may result in important details being overlooked. This paper explores methods of showing administrators quantitative data in meaningful ways so as to better look for anomalies or patterns from sources such as intrusion detection systems, port scanning tools and firewalls.

The paper shows interesting ways of converting vast quantities of event types into graphs and other visual representations. This included splitting the events into different types: Network traces, security events, network activity context, user/asset context, network events and application logs. The image to the right shows 80 hours of network data on a network of 1020 hosts. The internal network is represented by the grid on the left, and external servers by squares on the right, wit square size denoting the level of activity. This is not extremely obvious upon first look.

One thing the paper does not do is really discuss the requirements of visualisation compared to merely parsing and reformatting text data. While the image above does a good job of showing that large amounts of data may be transferred to external servers from internal sources, we have no easy way of knowing if any of this data is malicious or coming from potentially unwelcome sources. If this data were formatted as an easily formatted excel table for example, we could apply filters very easily to look for patterns in data. This is a common theme in the paper, where visualisations show what at first seems to be useful information, but allows little to no exploration of potential anomalies once identified short of digging through raw data.

##Citation

Shiravi, H., Shiravi, A., & Ghorbani, A. a. (2012). 006 A survey of visualization systems for network security. *IEEE Transactions on Visualization and Computer Graphics*, *18*(8), 1313–1329. https://doi.org/10.1109/TVCG.2011.144

- URL: <http://jamia.oxfordjournals.org/content/jaminfo/18/5/544.full.pdf>   
- Title: Natural Language Processing: An Introduction  
- Authors: Prakash M Nadkarni, Lucila Ohno-Machado, Wendy W Chapman  
- Notes: Aimed at a medical audience but could still be relevant as an NLP intro

##Review

This introduction to Natural Language Processing acts as an overview of common machine-learning approaches currently being used and possible future directions of NLP, as well as some of the associations with IR – Information Retrieval. One of the first things defined is that of statistical NLP –NLP based on machine learning methods, learning via large annotated bodies of text which provided the standard they were looking to achieve. The paper quickly becomes complex, looking into data driven approaches to NLP and their drawbacks such as Hidden Markov Models (HMMs), which is a system where variables can switch between several states and generate possible outputs. The issue with HMM’s is that we can only see the output, not the process at it takes to comes to that output.

While this paper is primarily looking into NLP, it more looks into the methods of machine learning that could be applied to NLP. However, it still lists and explains a good number of NLP sub-problems which, while primarily aimed toward the medical field, are still relevant in computer science, such as sentence boundary detection and morphological decomposition – The act of separating words into smaller words.

The paper also has an interesting section focussing on the future of artificial intelligence and NLP, quoting heavily IBM’s Watson supercomputer and its attempt at beating humans in the game Jeopardy. With 16TB ram, Watson is designed to hold all of reference content in memory, as opposed to being disk-I/O-bound, which makes its seek time exceptionally slow. However, Watson can be easily misled with certain questions – Asking it “Which US city has two airports, one named after a World War II Battle, the other after a World War II Hero?” would be a multi-step process which Watson could not answer, as the reference content used for machine learning was structured as one sentence question and answers (“What/who is/are X?”).

##Citation

Nadkarni, P. M., Ohno-Machado, L., Chapman, W. W., Manning, C., Raghavan, P., Schuetze, H., … Miller, R. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 544–51. https://doi.org/10.1136/amiajnl-2011-000464