ASwatch: An AS Reputation System to Expose Bulletproof Hosting ASes – Public Review

Walter Willinger
NIKSUN, Inc.
Princeton, NJ, USA
wwillinger@niksun.com

The authors have to be congratulated for writing a paper that prompted amazingly consistent and overall very positive reactions from all the reviewers. The reviewers expressed very similar sentiments about the work described in the paper and were all (i) excited about the problem, (ii) lukewarm about the methodology, and (iii) intrigued by the potential of the work.

To explain the reviewers' excitement, the paper is the first to propose an AS reputation system that relies exclusively on control-plane data. As such, their system ASwatch has the potential to be used pro-actively to defend against malicious ASes that exist for one main reason – to support various forms of cybercrime. Blocking their entire traffic is not only more effective and efficient than blocking traffic that is sent from selective prefixes or IPs associated with such malicious ASes, but the risk of collateral damage (i.e., blocking legitimate traffic) is largely non-existent when dealing with these so-called bulletproof hosting ASes.

While the reasons for the reviewers' lukewarm reaction to the proposed methodology vary, they all concern in one form or another a recent development in data-driven networking research where Machine Learning has become a "hammer" that makes every problem look like a "nail." Indeed, ASwatch illuminates this trend: Use domain knowledge to extract statistical features from available measurements (i.e., BGP-based data), build and train a statistical classifier based on ground truth (i.e., past known malicious ASes), and apply the obtained model to a new set of ASes to identify likely "bad" ASes among them. In the context of applying this popular recipe to the problem at hand, some of the issues raised by the reviewers included "Why these features and not others?", "Why this classifier and not some other?", "What impact does the small training set have on, say, over-training the chosen classifier?", "How useful is ASwatch in practice in view of the reported high false-positive rate?", "How does the subpar quality of some of the mined data impact the classification?", and "What (if any) new insights can be gained from this ML-based 'black box' approach?" The reviewers felt that a number of these and related issues remained under-explored and deserved more attention.

Finally, when judging the potential of the work, there was unanimous agreement among the reviewers that the design of an ASwatch-like system that uses control-plane data (as advocated in this paper) and combines it with data-plane measurements (as explored in earlier papers) would be a promising future direction for this work and would be a significant step towards deploying pro-active and effective defense mechanisms against ASes that exist exclusively for the purpose of supporting cybercrime. Recent publicity about a decision by Level 3 to start blocking traffic suspected of being the result of criminal activities (e.g., see http://www.wsj.com/articles/level-3-tries-to-waylay-hackers-1432891803) illuminates the enormous value that an ASwatch-like system could have in practice.

In summary, despite a certain amount of lack of enthusiasm for the chosen ML approach, the reviewers rallied around this paper for three main reasons. First, the paper addresses an interesting and important problem; second, the problem has been under-studied in the past and deserves to be more thoroughly explored; and third, the paper relies on an intriguing and original idea (i.e., use of control-plane data) and can be expected to generate significant follow-up work, though only time will tell!