- M2 算法:本文提出的基于词素特征的轻量级域名检测算法.在以词素为单元变长切分域名的基础上,通过统计 6 个词素特征测度,并应用现有的 C4.5 算法分类合法域名和恶意域名.
- M3 算法:为了说明以词素为单元的域名字面特征能够保留以单词为单元的域名语言学特征,这里将 M2 算法所依赖的词素库换成单词库,其余所有操作不变.

本文基于合法域名集 *Good_Domain_Set* 和恶意域名集 *Malicious_Domain_Set*,使用十字交叉验证法,从检测准确率、假阳性(合法域名检测为恶意域名)和假阴性(恶意域名检测为合法域名)等 3 个方面,评估并比较上述 3 种基于字面特征的轻量级域名检测算法.如图 9 所示.

- 基于二元组频率分布统计的 M1 算法具有最低的检测准确率 51.7%、最高的假阳性 43.2%和最低的假阴性 5.1%;进一步分析 M1 算法在两个样本集上各自的检测准确率,发现其对于恶意域名样本集的检测可以达到 89.8%,而对于合法域名样本集的检测仅为 13.7%;
- 基于单词特征的 M3 算法具有最高的检测准确率 71.7%、最低的假阳性 14.2%和较低的假阴性 14.1%,且该算法在两个样本集上表现出相似的准确率(前者 71.7%,后者 71.8%);
- 基于词素特征的 M2 算法,其三方面的评估结果(69.9%检测准确率、16.4%假阳性和 13.7%假阴性)稍逊于 M3 算法(准确率相对偏低 2.5%,假阳性相对偏高 15.5%,假阴性却相对偏低 2.8%),但明显优于 M1 算法(准确率相对高 35.2%,假阳性相对低 62.0%,假阴性却相对高 168.6%),且其对两个样本集也具有相近的检测准确率(合法域名样本集 67.3%,恶意域名样本集 72.5%).

综上所述:一方面,M2 算法具有与 M3 算法相似的检测准确率,证明以词素为单元的域名字面特征能够保留以单词为单元的域名语言学特征;另一方面,M2 算法比 M1 算法提高 1/3 的准确率,说明域名内含的词素特征比二元组频率分布特征更能刻画域名的语言学特征,可以更有效地作为特征测度用于分类合法域名和恶意域名.

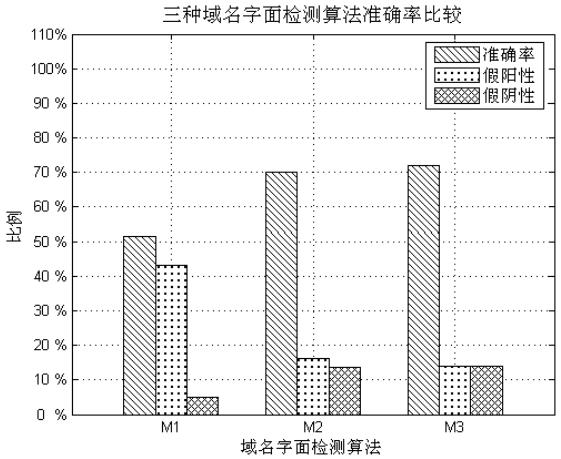



Fig.9 Detection accuracy assessment of the three algorithms

图 9 3 种域名字面检测算法的准确率评估

2.2.2 准确率影响因素

本文提出的 6 个分类测度,都是建立在一定数量的三层域名标签基础上,对其内部所含词素数目、长度等的数量统计.就数学统计方法本身而言,随机抽取的样本容量越大,用样本估计出的总体特性就越接近真实.因此,二层域名标签所辖的三层域名标签数作为关键因子,会直接影响检测算法最终的准确率.本文针对基于词素特征的轻量级域名检测算法 M2,分别选取所辖三层域名标签数在[5,10)范围内的 4 140 个二层域名标签对象、在[10,30)范围内的 2 800 个二层域名标签对象和在[30,+∞)范围内的 1 820 个二层域名标签对象,仍从检测准确率、假阳性和假阴性这 3 个方面,评估三层域名标签数对 M2 检测算法准确率的影响.

如图 10 所示:随着所辖三层域名标签数的增加,M2 算法的准确率有很大程度的上升,从 63.9%上升到

68.3%(相对上升 6.9%),再上升到 74.6%(相对上升 16.5%);假阳性和假阴性也有明显的下降,假阳性从 20.0%下降到 17.2%(相对下降 19.0%),再下降到 11.5%(相对下降 42.5%);假阴性也从 16.1%下降到 14.5%(相对下降 9.9%),再下降到 13.9%(相对下降 13.7%).

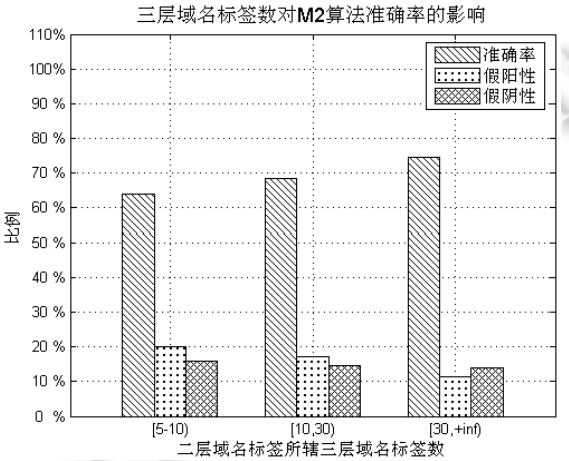



Fig.10 Influence of the third-level domain label number to the detection accuracy of algorithm M2

图 10 三层域名标签数对 M2 检测算法准确率的影响

虽然随着三层域名标签数的增加,M2 算法的准确率有显著的提高,但是恶意域名集中的样本数量也会显著地减少,从而影响抽样样本对总体的估计.本文权衡两方面,选取三层域名标签数超过 5 的二层域名标签集,作为算法介绍和评估比较时统一使用的样本空间.

### 2.3 抗逃避能力

传统的基于域名字符串长度、所含字母数目、数字数目等字面特征的词法分析方法,很容易被攻击者借助事前相应的特征统计来逃避.为此,本节以合法域名集 *Good_Domain_Set* 和恶意域名集 *Malicious_Domain_Set* 为训练集,以算法自动生成的随机域名集 *Random_Domain_Set*,*Dict_Domain_Set*,*Kwyjibo_Domain_Set* 和 *Pop_Domain_Set* 为测试集,从准确率角度评估比较上述 3 种算法的抗逃避能力.由于恶意域名样本集中缺少 *Dict_Domain_Set* 和 *Kwyjibo_Domain_Set* 中的随机域名样本,因此在恶意域名集中分别增加 400 组随机域名.此外,对恶意域名集的检测不存在假阳性,即,准确率和假阴性此消彼长,评估时只需关注准确率一个方面.如图 11 所示.

- 针对随机域名集 *Random_Domain_Set*(域名符合合法域名的二元组频率分布特征),M1 算法基本失效(准确率仅为 1.23%),但 M2 和 M3 算法基本上能完全检测(前者准确率 98.0%,后者 100%);
- 针对随机域名集 *Dict_Domain_Set*(域名直接使用英语单词构成,同时,在其尾部增加随机数字串),M1 和 M3 算法能够完全检测,M2 算法也具有 99.2%的准确率;
- 针对域名集 *Kwyjibo_Domain_Set*(使用 Kwyjibo 工具生成的形似单词的域名,同时,也在其尾部增加随机数字串),M1 算法能够完全检测,M2 和 M3 算法也基本上能完全检测(前者准确率 99.4%,后者 99.1%);
- 而对于 *Pop_Domain_Set*(替换知名域名的二层标签,保留或重构其三层标签),M2 和 M3 算法都具有较低的准确率(前者 32.8%,后者 28.3%),但是 M1 算法由于误报较高,反而拥有 86.3%的检测准确率.

综上所述,基于二元组频率分布统计的检测算法,在面对攻击者事前经过二元组频率分布统计后生成的域名时基本失效.而本文提出的基于词素特征的检测算法,能够同时抗拒攻击者通过事前特征统计的逃避策略以及借助字典或 Kwyjibo 工具的随机域名生成策略.但是对于重用知名域名三层标签的逃避策略,基于词素特征和单词特征的检测算法都表现出较低的检测能力.

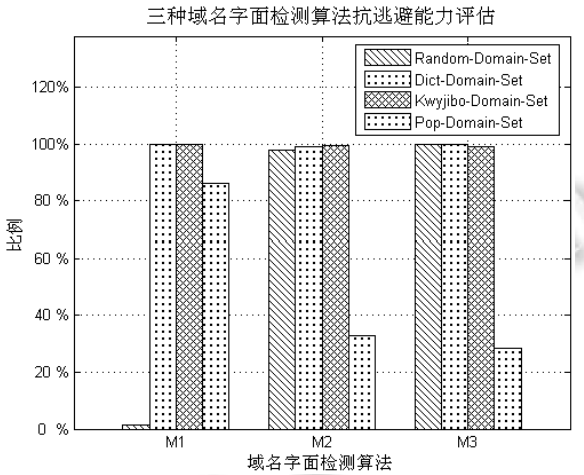



Fig.11 Anti-interference ability assessment of the three algorithms

图 11 3 种域名字面检测算法的抗逃避能力评估

### 2.4 系统开销

作为轻量级域名检测算法,首先需要保证其具有较低的内存开销和计算复杂度,以便能在有限的系统资源和计算时间内尽可能多地检测出可疑域名.本节选取合法域名样本集 *Good_Domain_Set* 和恶意域名样本集 *Malicious_Domain_Set* 作为训练集,在普通 PC 机上(Intel(R) Xeon(R)单核 cpu,频率 2.00GHz,内存 2G,Linux 系统版本 2.6.15-23-386),分别使用上述 3 种域名字面检测算法对实测域名集 *JS_Domain_Set* 进行检测,并在此过程中,从理论和实际两个角度分析其运行所需的内存和时间开销.

表 3 中,

- 基于二元组频率分布统计的 M1 算法具有最低的内存开销和计算复杂度;
- 与 M1 算法相比,基于词素特征的 M2 算法则具有相对较大的内存开销(临时内存空间增加 27.82MB,相对增长 53.2%,常驻内存空间相对增加 2.69MB)和计算复杂度(理论计算操作次数增加 66.5 倍,实际运行时间增长 21.5 倍);
- 而基于单词特征的 M3 算法,其与 M2 算法相比具有相同的临时内存开销,常驻内存增加 2.25MB,相对增长 83.0%,理论计算复杂度增加 1.8 倍,实际运行时间增长 33.0%.

**Table 3** Space/Time complexity comparison between the three algorithms

**表 3** 3 种算法内存开销和计算复杂度比较

| 算法 | 内存开销 | 计算复杂度 |
|---|---|---|
| M1 | 临时 52.32MB,常驻 0.02MB | 理论 111.3M 次操作,实际 8.7s |
| M2 | 临时 80.14MB,常驻 2.71MB | 理论 7 508.6M 次操作,实际 196.1s |
| M3 | 临时 80.14MB,常驻 4.96MB | 理论 21 333M 次操作,实际 260.8s |

### 2.5 实用测试

运用本文提出的基于词素特征的域名字面检测算法 M2,通过合法域名集 *Good_Domain_Set* 和恶意域名集 *Malicious_Domain_Set* 的训练学习,对从中国教育科研网江苏省网边界实际采集的域名集 *JS_Domain_Set* 进行检测,共发现 745 个可疑二层域名标签及其子域名.进一步分析发现,其中 199 个二层域名所辖子域名中含有黑名单中出现过的恶意域名,13 个二层域名未经注册,118 个二层域名包含色情、赌博、虚假婚介、恶意销售等内容,65 个二层域名所辖网站无效、过期或筹建中,58 个二层域名包含合法的政府、学校和公司网站,另外 292 个二层域名无法通过网站直接访问.综上所述,除去无法确认的 292 个二层域名,剩余 453 个二层域名中,能够确认

395 个为恶意二层域名,58 个为合法二层域名,即:基于词素特征的域名字面检测算法的实际检测准确率为87.2%,假阳性为 12.8%.

## 3 总 结

网络安全监测需要在最短的时间内尽可能多地检测出可疑域名.面对网络中实际使用的庞大域名对象,传统基于 DNS 交互报文的 DPI 检测技术由于资源开销过大,难以满足现实的性能需求.

本文基于域名自身字面特征,提出一种轻量级的检测算法,能够快速感知和标识恶意服务使用的可疑域名,以便有针对性地使用现有的更为复杂和更为准确的算法.该轻量级域名检测算法选取自然语言中最小的语义单元词素,设计启发式字符串切割算法来快速挖掘域名中蕴含的语言学特征,并在二层域名标签聚类的基础上,提出一组基于词素特征的检测测度,用于 C4.5 算法以实现合法域名和恶意域名的分类.

实验结果表明:本文提出的词素特征比 $n$ 元组频率分布特征具有更高的检测准确率(准确率相对增长35.2%,假阳性相对降低 62.0%),且能够有效地抵挡攻击者借助事前相应特征统计的逃避策略(几乎能够完全检测符合合法域名二元组频率分布特征的域名以及借助字典或者 Kwyjibo 工具自动产生的随机域名).但是在面对重用知名域名三层标签的逃避策略时,表现出较低的检测能力,准确率只有 32.8%.进一步应用该算法对中国教育科研网江苏省网边界实际采集到的域名集进行检测,实测结果表明,该算法具有较高的检测准确率(87.2%)、较低的内存开销(80.14MB 的临时内存,2.71MB 的常驻内存开销)和计算复杂度(运行时间 196.1s).此外,本文还比较了基于词素和基于单词的两种字面特征检测方法,单词特征虽然具有略高的检测准确率(准确率相对偏高 2.7%,假阳性相对偏低 13.8%,假阴性却相对偏高 2.7%),但是其常驻内存开销和计算复杂度均明显高于词素特征(常驻内存空间相对增加 83.0%,理论计算复杂度相对增加 1.8 倍).

因此,本文提出的基于词素特征的域名字面检测算法能够替代现有基于 $n$ 元组频率分布特征和基于单词特征的检测算法,同时满足轻量级算法对系统开销和检测准确率的需求.

**References**:

[1] Porras P, Saidi H, Yegneswaran V. A foray into Conficker's logic and rendezvous points. In: Lee W, ed. Proc. of the 2nd USENIX Conf. on Large-Scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More (LEET 2009). Boston: USENIX, 2009.

[2] Conficker C Analysis. 2009. http://mtc.sri.com/Conficker/addendumC

[3] Royal P. Analysis of the Kraken Botnet. 2008. https://www.damballa.com/downloads/r_pubs/KrakenWhitepaper.pdf

[4] Stone-Gross B, Cova M, Cavallaro L. Your botnet is my botnet: analysis of a botnet takeover. In: Al-Shaer E, Jha S, Keromytis AD, eds. Proc. of the 16th ACM Conf. on Computer and Communications Security (CCS 2009). Chicago: ACM Press, 2009. 635−647. [doi: 10.1145/1653662.1653738]

[5] Chatzis N, Popescu-Zeletin R. Flow level data mining of DNS query streams for email worm detection. In: Corchado E, Zunino R, Gastaldo P, Herrero A, eds. Proc. of the Int'l Workshop on Computational Intelligence in Security for Information Systems (CISIS 2008). Berlin, Heidelberg: Springer-Verlag, 2009. 186−194. [doi: 10.1007/978-3-540-88181-0_24]

[6] Chatzis N, Popescu-Zeletin R. Detection of email worm-infected machines on the local name servers using time series analysis. Journal of Information Assurance and Security, 2009,4(3):292−300.

[7] Chatzis N, Popescu-Zeletin R, Brownlee N. Email worm detection by wavelet analysis of DNS query streams. In: Dasgupta D, Zhan J, eds. Proc. of the IEEE Symp. on Computational Intelligence in Cyber Security (CICS 2009). Nashville: IEEE, 2009. 53−60. [doi: 10.1109/CICYBS.2009.4925090]

[8] Chatzis N, Brownlee N. Similarity search over DNS query streams for email worm detection. In: Awan I, ed. Proc. of the 2009 Int'l Conf. on Advanced Information Networking and Applications (AINA 2009). Bradford: IEEE, 2009. 588−595. [doi: 10.1109/AINA.2009.132]

[9] Caglayan A, Toothaker M, Drapeau D, Burke D, Eaton G. Real-Time detection of fast flux service networks. In: Walter E, ed. Proc. of the 2009 Cybersecurity Applications & Technology Conf. for Homeland Security (CATCH 2009). Washington: IEEE, 2009. 285−292. [doi: 10.1109/CATCH.2009.44]

[10] Choi H, Lee H, Kim H. Botnet detection by monitoring group activities in DNS traffic. In: Wei D, ed. Proc. of the 7th IEEE Int'l Conf. on Computer and Information Technology (CIT 2007). Fukushima: IEEE, 2007. 715−720.

[11] Choi H, Lee H, Kim H. BotGAD: Detecting botnets by capturing group activities in network traffic. In: Bosch J, Clarke S, eds. Proc. of the 4th Int'l ICST Conf. on Communication System Software and Middleware (COMSWARE 2009). Dublin: ACM Press, 2009. [doi: 10.1145/1621890.1621893]

[12] Choi H, Lee H. Identifying botnets by capturing group activities in DNS traffic. Computer Networks: The Int'l Journal of Computer and Telecommunications Networking, 2012,56(1):20−33. [doi: 10.1016/j.comnet.2011.07.018]

[13] Antonakakis M, Perdisci R, Lee W, Vasiloglou N, Dagon D. Detecting malware domains at the upper DNS hierarchy. In: Wagner D, ed. Proc. of the 20th USENIX Conf. on Security (SEC 2011). San Francisco: USENIX, 2011.

[14] Antonakakis M, Perdisci R, Nadji Y, Vasiloglou N, Abu-Nimeh S, Lee W, Dagon D. From throw-away traffic to bots: Detecting the rise of DGA-based malware. In: Kohno T, ed. Proc. of the 21st USENIX Conf. on Security Symp. (Security 2012). Bellevue: USENIX, 2012. 491−506.

[15] Bilge L, Sen S, Balzarotti D, Kirda E, Kruegel C. Exposure: A passive DNS analysis service to detect and report malicious domains. ACM Trans. on Information and System Security (TISSEC), 2014,16(4). [doi: 10.1145/2584679]

[16] Ma J, Saul LK, Savage S, Voelker GM. Beyond blacklists: Learning to detect malicious web sites from suspicious URLs. In: Elder J, Fogelman FS, Flach P, Zaki M, eds. Proc. of the 15th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2009). Paris: ACM Press, 2009. 1245−1254. [doi: 10.1145/1557019.1557153]

[17] Ma J, Saul LK, Savage S, Voelker GM. Learning to detect malicious URLs. ACM Trans. on Intelligent Systems and Technology (TIST), 2011,2(3):493−500. [doi: 10.1145/1961189.1961202]

[18] Prakash P, Kumar M, Kompella RR, Gupta M. PhishNet: Predictive blacklisting to detect phishing attacks. In: Mandyam G, Westphal C, eds. Proc. of the 29th Conf. on Information Communications (INFOCOM 2010). San Diego: IEEE, 2010. 346−350. [doi: 10.1109/INFCOM.2010.5462216]

[19] Yadav S, Reddy AKK, Reddy ALN, Ranjan S. Detecting algorithmically generated malicious domain names. In: Allman M, ed. Proc. of the 10th ACM SIGCOMM Conf. on Internet Measurement (IMC 2010). Melbourne: ACM Press, 2010. 48−61. [doi: 10.1145/1879141.1879148]

[20] Yadav S, Reddy AKK, Reddy ALN, Ranjan S. Detecting algorithmically generated domain-flux attacks with DNS traffic analysis. IEEE/ACM Trans. on Networking (TON), 2012,20(5):1663−1677. [doi: 10.1109/TNET.2012.2184552]

[21] Khaitan S, Das A, Gain S, Sampath A. Data-Driven compound splitting method for English compounds in domain names. In: Cheung D, Song IY, Chu W, Hu XH, Lin J, eds. Proc. of the 18th ACM Conf. on Information and Knowledge Management (CIKM 2009). Hong Kong: ACM Press, 2009. 207−214. [doi: 10.1145/1645953.1645982]

[22] Srinivasan S, Bhattacharya S, Chakraborty R. Segmenting Web-domains and hashtags using length specific models. In: Chen XW, Lebanon G, Wang HX, Zaki MJ, eds. Proc. of the 21st ACM Int'l Conf. on Information and Knowledge Management (CIKM 2012). Maui Hawaii: ACM Press, 2012. 1113−1122. [doi: 10.1145/2396761.2398410]

[23] Marchal S, Francois J, State R, Engel T. Proactive discovery of phishing related domain names. In: Stolfo SJ, Stavrou A, Wright CV, eds. Proc. of the Research in Attacks, Intrusions, and Defenses. Berlin, Heidelberg: Springer-Verlag, 2012. 190−209. [doi: 10.1007/978-3-642-33338-5_10]

[24] Schiavoni S, Maggi F, Cavallaro L, Zanero S. Tracking and characterizing botnets using automatically generated domains. CoRR, 2013. http://arxiv.org/pdf/1311.5612.pdf

[25] Plag I. Word-Formation in English. Cambridge: Cambridge University Press, 2002.

[26] Alexa. 2014. http://www.alexa.com/topsites/

[27] Palevo tracker. 2014. https://palevotracker.abuse.ch/

[28] Zeus tracker. 2014. https://zeustracker.abuse.ch/

[29] DNS-BH—Malware domain blocklist. 2014. http://www.malwaredomains.com/

[30] Malware domain list. 2009. http://www.malwaredomainlist.com

[31] PhishTank. 2014. http://www.phishtank.com/

[32] Blacklist provided by joewein.net (JWSDB). 2014. http://joewein.net/spam/blacklist.htm

[33] Baddeley A, Della Sala S. Working memory and executive control. Philosophical Trans. of the Royal Society of London Series B—Biological Sciences, 1996,351(1346):1397−1403.

[34] Kotsiantis SB. Supervised machine learning: A review of classification techniques. In: Maglogiannis I, Karpouzis K, Wallace M, Soldatos J, eds. Proc. of the 2007 Conf. on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies. Amsterdam: IOS Press, 2007. 3−24.

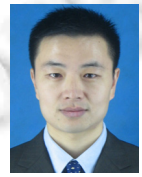
[35] Crawford H, Aycock J. Kwyjibo: Automatic domain name generation. Software Practice and Experience, 2008,38(14):1561−1567. [doi: 10.1002/spe.885]

[36] Quinlan JR. C4.5: Programs for Machine Learning. San Francisco: Morgan Kaufmann Publishers Inc., 1993.

**张维维**(1984－),男,江苏南通人,博士生,主要研究领域为网络安全.

**龚俭**(1957－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为网络管理,网络安全.

**刘茜**(1991－),女,硕士生,主要研究领域为网络安全.



**刘尚东**(1979－),男,博士生,CCF 会员,主要研究领域为网络安全.

**胡晓艳**(1985－),女,博士生,主要研究领域为网络管理,下一代互联网.