Attacklets: Modeling High Dimensionality in Real World Cyberattacks

Cuneyt G. Akcora*, Jonathan Z. Bakdash† ‡, Yulia R. Gel*, Murat Kantarcioglu*,
Laura R. Marusich§, Bhavani Thuraisingham*
*University of Texas at Dallas
{cuneyt.akcora, bhavani.thuraisingham, ygl, muratk}@utdallas.edu
†U.S. Army Research Laboratory South at the University of Texas at Dallas, ‡Texas A&M–Commerce
jonathan.z.bakdash.civ@mail.mil
§U.S. Army Research Laboratory South at the University of Texas at Arlington, laura.r.marusich.ctr@mail.mil

***Abstract*—We introduce attacklets, a novel approach to model the high dimensional interactions in cyberattacks. Attacklets are implemented using a real-world dataset of cyberattacks from the Verizon Data Breach Investigation Report. Whereas the commonly used attack graphs model the action sequences of attackers for specific exploits, attacklets model general attributes and states of each attack separately. Attacklets may inform the number and types of attributes across a wide range of cyberattacks. These structural properties can then be used in machine learning models to classify and predict future cyberattacks.**

***Index Terms*—cybersecurity; attack graph; attacklet; cyberattack; data depth**

I. Introduction

A common quantitative method for modeling cyberattacks uses attack graphs, or attack trees, which are a graphical representation of attacker actions and corresponding system security states [1]. In contrast, many methods for modeling risks in cybersecurity are qualitative because they use subjective data such as expert opinion [2]. Potential applications for attack graphs include determining overall risks in the network and identifying and mitigating specific exploits and vulnerabilities [1], [3], [4]. Here, we demonstrate a novel approach to modeling cyberattacks by leveraging the high dimensionality and subdimensionality of attack attributes. We call this approach **attacklets**. Attacklets are implemented using a real–world dataset of cyberattacks from the Verizon Data Breach Investigations Report (DBIR) [5]. Compared to typical attack graphs, attacklets model the high dimensional interactions using the data depth for five attributes/states of co-occuring attack attributes and their nested sub-attributes.

A. Background and Motivation

The DBIR dataset provides a unique attack graph because of its high dimensionality (see II). Existing attack graphs either evaluate every attack path or use assumptions to simplify the graph [6], [7]. A common attack graph is state-based, and consists of all possible security states, which may not be computationally feasible. Computational challenges are typically addressed by limiting the number of states or applying techniques for minimization analysis [7], [8].

An alternative approach of addressing computational issues for state-based attack graphs is to summarize all attack paths using descriptive statistics, such as the mean and standard deviation of path lengths and fixed values for the popularity of specific exploits [9]. One notable exception, a computationally feasible state-based attack graph uses approximate inference of all attack paths [6]. Another type of attack graph simplifies the assumptions between attacks and system security states (e.g., necessary and sufficient preconditions for attack success, security states can only stay the same or get worse) [6], [10]. Whereas, due to individual processing of each attack, our attacklet approach is both computationally efficient and allows for longitudinal analyses with attacklet summaries.

II. Dataset

The DBIR dataset is publicly available from [11], see [5] for a detailed description. The dataset contains breaches from 1971 to 2017. The dataset schema is the Vocabulary for Event Recording and Incident Sharing (VERIS) [12], which contains 24 top-level variables such as *impact*, *reference*, and *summary*. Using VERIS, we focus on common attributes of the breaches: (**Discovered_by**) which gives the discovery source of a breach and the four main attributes: **Actor, Action, Attribute,** and **Asset**. These attributes are chosen to provide a succinct and informative representation of each breach. We refer to this reduced dataset as the $\mathbb{A}^4\mathbb{D}$ model. Each attribute is further sub-coded with a subset of states. The Discovery attribute has *external*, *other*, *partner*, *unknown* and *internal* states to code the source of attack discovery. The Action that was taken in the attack is sub-coded with *error*, *hacking*, *malware*, *misuse*, *physical*, *environmental*, *social* and *unknown*. The Actor attribute has the states *external*, *internal*, *partner* and *unknown* to code the origin of the attack. Target of the attack is coded with the Asset attribute, which has *media*, *network*, *people*, *server*, *unknown*, and *user* states. Attribute encodes the result of the attack with loss of asset *availability*, breach of asset *confidentiality*, and violation of asset *integrity*. In

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. ARL or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation. This work was supported by the U.S. ARL Senior Fellowship Program (L.R.M.) administered by the Oak Ridge Associated Universities under Cooperative Agreement Number W911NF-17-2-0003, and by NSF IIS 1633331, NSF DMS 1736368, NSF ECCS 1824716 and ARO W911NF-17-1-0356. 978-1-5386-7848-0/18/$31.00 ©2018 IEEE.

the dataset, 7,614 (i.e., 99.7%) of breaches contain the $\mathbb{A}^4\mathbb{D}$ attributes. Breach reports tend to contain one state (median is 1) in each attribute (e.g., Actor: external), but reports with multiple states for the same attribute also exist (e.g., Action: misuse and error). In maximum, breach reports contain Action: 5, Actor: 4, Asset: 5, Attribute: 3 and Discovery: 1 states.

III. The Attacklet Model

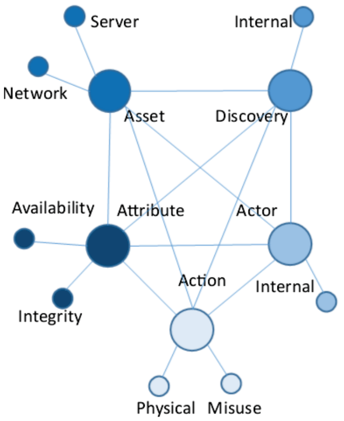
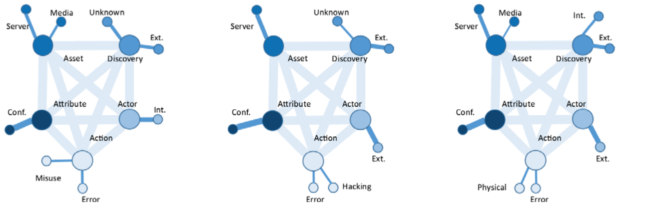


Fig. 1: The attacklet of a report where an internal worker's physical actions and misuse caused the integrity of server and network assets to be violated, making the assets unavailable for use.

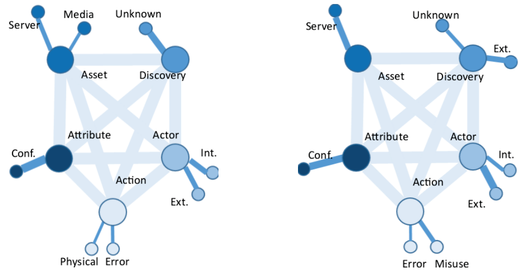
We use the term **attacklet** to refer to the attack graph created from a breach report where nodes are $\mathbb{A}^4\mathbb{D}$ attributes ($\mathcal{X}$) and their states ($\mathcal{S}$). Each attacklet edge $e = (u, x)$ where $u, v \in \mathcal{X} \cup \mathcal{S}$. Attacklets can be visualized as in Figure 1; as they co-appear in the breach report, all $\mathbb{A}^4\mathbb{D}$ nodes are connected. A state is connected to its attribute, such as Internal in the Discovery attribute. For each attribute, each state has an associated position on the graph. Next, we will use these positions to summarize the information contained in a set of attacklets.

Attacklet Summaries: Attacklets can be grouped or analyzed in two additional aspects: **country** of the victim and **year** of the attacklet. The year attribute allows a longitudinal study of how attacklets change shape through years. In Figure 2, we summarize the attacklets for the last three years of the dataset (i.e., 2015–2017) with a 50% resolution, which visualizes the top 50% most frequent states only, for a better view. Annual summaries show that most breaches occur due to error actions, and servers are the most common target in confidentiality attacks. We create a similar view for the attacklets originating from USA (78% of all attacklets) vs. other countries (i.e., Non-USA). Figure 3 shows that USA attacklets differ from the Non-USA ones in Action, Asset and Discovery attributes. For example, the USA attacklets show that there are more media assets attacked, and more physical actions than in the the Non-USA counterparts.



(a) 2015 (b) 2016 (c) 2017

Fig. 2: Summary attacklets for the most recent three years (50% resolution).Edge thickness indicates the percentage of an edge appearing in all attacklets.



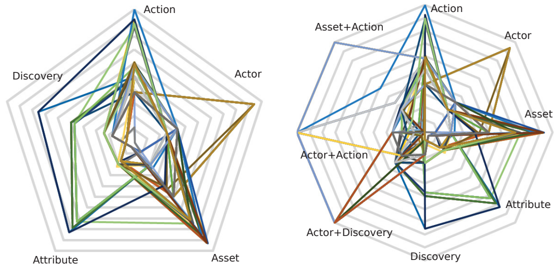
(a) USA (b) Non-USA

Fig. 3: Attacklet summaries for USA vs. other countries (50% resolution). Edge thickness indicates the percentage of an edge appearing in all attacklets.

IV. Attacklet Representation

A key goal of this work is to create a representation of attacklets so that tasks, such as attacklet classification and clustering, can be carried out for security applications (e.g., anomaly detection). To evaluate properties of attacklet distributions, we start with selecting a *reference*, or a *baseline* attacklet which can serve multiple purposes. First, a reference attacklet can be viewed as the most representative type of breach in a given year or geographic region. We can then assess how different countries differ from this baseline in terms of the most typical breaches. Second, we can study temporal deviations of the newly incoming attacklets from the earlier defined reference attacklet, which, in turn, can be used as a early warning signal for new types of breaches. Third, a reference attacklet can be used for visualization.

In the most basic model, an attacklet can be represented with respect to the reference attacklet in five $\mathbb{A}^4\mathbb{D}$ attributes. Distance of an attacklet in each attribute can be computed



(a) Attacklets in $\mathbb{A}^4\mathbb{D}$ attributes. (b) Attacklets in $\mathbb{A}^4\mathbb{D}$ and three pairwise attributes.

Fig. 4: [Color online] A representation of the latest 100 attacklets from 2017 in the dataset. Distance (0 at the center, 1 at the outer rim) in each of the $\mathbb{A}^4\mathbb{D}$ attributes + 3 pairwise attributes indicates deviation from the reference attacklet. An attacklet closes a planar surface with its distance to the reference attacklet in multiple attributes.

in isolation. The reference attacklet of the attribute becomes a vector of attribute state probabilities. We compute the distance of attacklet from the reference attacklet as $1 - sim(\text{attacklet}, \text{reference})$ where the sim function is the Cosine similarity [13]. Figure 4a shows attacklets in a radar plot where we associate five distance values for an attacklet.

In a second approach, we consider more than the five attributes using a joint multivariate distribution of attributes. Formally, let $\boldsymbol{\mathcal{X}}_i, i = 1, \ldots, 5$ be a $\mathbb{A}^4\mathbb{D}$ attribute (e.g., Asset). Each $\boldsymbol{\mathcal{X}}_i, i = 1, \ldots, 5$ can be viewed as a nominal (categorical) discrete random variable which can attain n_i different states (e.g., Asset: server), and each state $\mathcal{S}_{ij}, j = 1, \ldots, n_i$ is coded as $1, \ldots, n_i$. For coding purposes, we sort the states $\mathcal{S}_{ij}$ in the alphabetic order; however, it is important to emphasize that $\mathcal{S}_{ij}$ is intrinsically qualitative and there exists no natural meaning behind ordering the states. With these notations, we can assess a marginal distribution P_i of the attribute $\boldsymbol{\mathcal{X}}_i$, where a probability of $\boldsymbol{\mathcal{X}}_i$ attaining the state $\mathcal{S}_{ij}$ is $\pi_j = |\mathcal{S}_{ij}| \,/\, |\boldsymbol{\mathcal{X}}_i|$, where states $j = 1, \ldots, n_i$. Here $|\mathcal{S}_{ij}|$ is the number of occurrences of the state $\mathcal{S}_{ij}$ of the attribute $\boldsymbol{\mathcal{X}}_i$ and N_i denotes the number of occurrences of the the attribute $\boldsymbol{\mathcal{X}}_i$. Multiple approaches exist to define a reference attacklet for the attribute $\boldsymbol{\mathcal{X}}_i$. Since $\boldsymbol{\mathcal{X}}_i$ is a categorical random variable, mode of a marginal distribution P_i may be one option, with the benefit that mode has an intrinsic natural meaning. However, extending the mode approach to analysis of multiple attributes is challenging since finding a mode of a multivariate distribution may have no analytic solution. As a result, the mode approach is largely limited to treating all attributes independently and does not account for dependencies among some key breach characteristics, which clearly play a significant role in assessing dynamics of breaches. For example, Actor: Internal and Discovery: Internal has a strong dependency, because breaches by internal actors may be noticed more often by internal resources.

$$\begin{aligned}&\mathbf{P}\{\boldsymbol{\mathcal{X}}_1; \ldots; \boldsymbol{\mathcal{X}}_5\} = \\ &P\{\boldsymbol{\mathcal{X}}_1 = \{\mathcal{S}_{11}, \ldots, \mathcal{S}_{1n_1}\}; \ldots; \boldsymbol{\mathcal{X}}_5 = \{\mathcal{S}_{51}, \ldots, \mathcal{S}_{5n_5}\}\}\end{aligned} \tag{1}$$

Since our primary focus is on evaluating relative characteristics of attacklets with respect to a reference attacklet, we define it, $\boldsymbol{\Upsilon}$, as an empirical mean of $\mathbf{P}$.

$$\begin{aligned}\boldsymbol{\Upsilon} = \sum_{\boldsymbol{\mathcal{X}}_1}\sum_{\boldsymbol{\mathcal{X}}_2}\sum_{\boldsymbol{\mathcal{X}}_3}\sum_{\boldsymbol{\mathcal{X}}_4}\sum_{\boldsymbol{\mathcal{X}}_5} \pi_{k,l,m,s,r}\big(&\boldsymbol{\mathcal{X}}_1 = k; \boldsymbol{\mathcal{X}}_2 = l; \\ &\boldsymbol{\mathcal{X}}_3 = m; \boldsymbol{\mathcal{X}}_4 = s; \boldsymbol{\mathcal{X}}_5 = r\big)\end{aligned} \tag{2}$$

where $\pi_{k,l,m,s,r} = \mathbf{P}\big(\boldsymbol{\mathcal{X}}_1 = k; \boldsymbol{\mathcal{X}}_2 = l; \boldsymbol{\mathcal{X}}_3 = m; \boldsymbol{\mathcal{X}}_4 = s; \boldsymbol{\mathcal{X}}_5 = r\big)$ is the probability to observe to observe an attacklet with the attributes $\big(\boldsymbol{\mathcal{X}}_1 = k; \boldsymbol{\mathcal{X}}_2 = l; \boldsymbol{\mathcal{X}}_3 = m; \boldsymbol{\mathcal{X}}_4 = s; \boldsymbol{\mathcal{X}}_5 = r\big)$ and is defined as the fraction of occurrences of the combination $\big(\boldsymbol{\mathcal{X}}_1 = k; \boldsymbol{\mathcal{X}}_2 = l; \boldsymbol{\mathcal{X}}_3 = m; \boldsymbol{\mathcal{X}}_4 = s; \boldsymbol{\mathcal{X}}_5 = r\big)$. Notice that a marginal distribution P_i is a univariate subcase of $\mathbf{P}$. Furthermore, (1) allows us to project $\mathbf{P}$ onto a subset of attributes which are of particular interest from a domain knowledge perspective. This knowledge can be as simple as 'an internal action such as misuse will more likely be discovered by other internal actors', which indicates a dependency between Action and Discovery. To illustrate the utility of pairwise attributes, we chose to use Discovery+Actor, Actor+Action and Action+Asset pairwise dependencies. Figure 4b shows attacklets with pairwise attributes.

While in this pilot study for simplicity we focus on a reference attacklet defined by $\boldsymbol{\Upsilon}$, a more general alternative is to employ a notion of data depth [14] which enables a more flexible multi-perspective evaluation of a multivariate distribution $\mathbf{P}$. With data depth [14], we can define multivariate quantiles of $\mathbf{P}$ (e.g., multivariate median as a reference attacklet, or the 5th and 95th multivariate percentiles of $\mathbf{P}$) which can be used to assign an anomaly score to an incoming attacklet.

V. Discussion and Conclusion

We have proposed attacklets – a novel, flexible and computationally efficient approach for modeling cyberattacks. We have illustrated utility of attacklet summaries for visualization and description of cyberattacks in different geographical regions and across years. Furthermore, we have developed a representation model to store attacklets and use them as input to machine learning models. In the future, we plan to develop data depth based anomaly detection algorithms on attacklets.

References

[1] O. Sheyner, J. Haines, S. Jha, R. Lippmann, and J. M. Wing, "Automated generation and analysis of attack graphs," in *Proceedings of IEEE Symposium on Security and Privacy*, 2002, pp. 1–12.

[2] L. Allodi and F. Massacci, "Security events and vulnerability data for cybersecurity risk estimation," *Risk Analysis*, vol. 37, no. 8, pp. 1606–1627, 2017.

[3] N. Poolsappasit, R. Dewri, and I. Ray, "Dynamic security risk management using bayesian attack graphs," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 1, pp. 61–74, 2012.

[4] R. P. Lippmann and K. W. Ingols, "An annotated review of past papers on attack graphs," Tech. Rep., 2005, retrieved August 13, 2018 from http://www.dtic.mil/dtic/tr/fulltext/u2/a431826.pdf.

[5] S. Widup, M. Spitler, D. Hylender, and G. Bassett, "2018 Verizon Data Breach Investigations Report 11th edition," Tech. Rep., Apr. 2018, https://www.verizonenterprise.com/verizon-insights-lab/dbir/.

[6] D. Sgandurra, A. Paudice, and E. C. Lupu, "Efficient attack graph analysis through approximate inference," *ACM Transactions on Privacy and Security (TOPS)*, vol. 20, no. 3, pp. 1–30, 2017.

[7] S. Jha, O. Sheyner, and J. Wing, "Two formal analyses of attack graphs," in *Proceedings of the 15th Annual IEEE Computer Security Foundations Workshop*, 2002, pp. 49–63.

[8] L. Wang, S. Noel, and S. Jajodia, "Minimum-cost network hardening using attack graphs," *ACM Computer Communications*, vol. 29, no. 18, pp. 3812–3824, 2006.

[9] N. Idika and B. Bhargava, "Extending attack graph-based security metrics and aggregating their application," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 1, pp. 75–85, 2012.

[10] P. Ammann, D. Wijesekera, and S. Kaushik, "Scalable, graph-based network vulnerability analysis," in *Proceedings of the 9th ACM Conference on Computer and Communications Security*, 2002, pp. 217–224.

[11] "VCDB: VERIS Community Database," retrieved Aug 13, 2018 from https://github.com/vz-risk/VCDB.

[12] "VERIS: Vocabulary for Event Recording and Incident Sharing (VERIS)," retrieved Aug 13, 2018 from https://github.com/vz-risk/veris.

[13] M. Steinbach, G. Karypis, V. Kumar *et al.*, "A comparison of document clustering techniques," in *Knowledge Discovery in Databases Text Mining (KDD TM) Workshop*, 2000, pp. 525–526.

[14] R. Y. Liu, "On a notion of data depth based on random simplices," *The Annals of Statistics*, pp. 405–414, 1990.