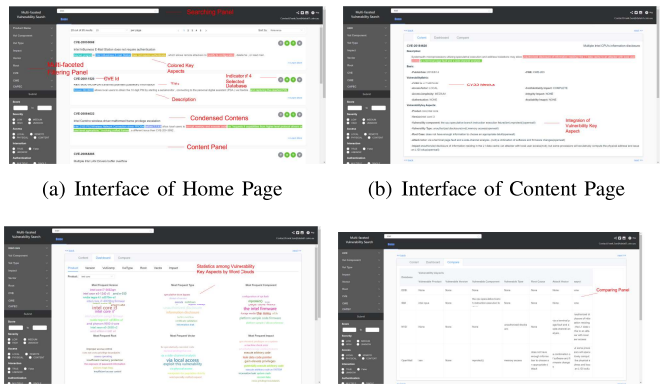
APPENDIX

A. Knowledge Base

We use online parser tools including ElementTree[2] and NLP toolkits including Spacy[3] to extract and clean the entities and relations. In total, we extract 572,816 entities (including 113,148 CVEs, 946 CWEs, 543 CAPECs, 50,039 product names, 207,054 versions, 78,550 vulnerable components, 7,069 vulnerability types, 27,759 root causes, 40,985 attack vectors, 46,703 impacts, and 20 elements from CVSS 2.0 metrics) and 2,868,473 relations except synonyms (including 1,687 *parentOf*, 1,687 *childOf*, 157 *canPrecede*, 157 *canFollow*, 2,364 *targetOf*, 232 *peerOf*, 113,148 *instanceOf*, 860 *sematicOf*, 221,325 *productOf*, 864,930 *versionOf*, 125,945 *componentOf*, 129,043 *typeOf*, 92,942 *rootOf*, 158,292 *vectorOf*, 274,904 *impactOf*, 146,800 *accessVectorOf*, 146,800 *complexityOf*, 146,800 *authenticationOf*, 146,800 *confidentialityOf*, 146,800 *integrityOf*, and 146,800 *availabilityOf*). Using patterns proposed by us, we extract 635,997,298 synonyms relations (including 5,166,822 synonyms relations among vulnerability type, 53,372,340 synonyms relations among root cause, and 577,458,136 synonyms relations among impact).

B. Website Implementation and Usage



(a) Interface of Home Page (b) Interface of Content Page

(c) Interface of Dashboard Page (d) Interface of Comparing Table Page

Fig. 4: Different Interfaces of Multi-faceted Vulnerability Searching Website

We implement a multi-faceted searching website based on our Vulnerability KG, whose UIs are shown in Fig. 4. The font end is achieved by Vue 2.0[4], and the back end is achieved by Django[5]. Fig. 4(a) shows the home page of the website. On the searching panel, users can use either CVE ids or key aspects to search the related CVEs. On the left multi-faceted filtering panel, users can filter searching results by characteristics including CWE, CAPEC, 7 vulnerability key aspects and CVSS metrics. On the content panel, we return the searching results with condensed contents, including titles, vulnerability descriptions, coloured key aspects and indicators of external links of 4 selected databases, which can provide user a quick look about the results. Users can click learn more button to transfer to content page in Fig. 4(b), which shows the detailed information of vulnerability characteristics.

By clicking Dashboard in Fig. 4(b), users will jump into dashboard panel in Fig. 4(c). On the dashboard panel, we provide statistics among vulnerability aspects, i.e., the most frequent vulnerable version of specific product. Users can have a direct visualization of the riskiest versions of specific products or the common impacts of specific impacts, which can help users to make prevention strategies based on possible vulnerable components, root causes, attack vectors and the impacts. Besides, we also provide statistics of historical CVSS 2.0 metrics of specific vulnerability key aspects offering reference of severity to the users.

By clicking Compare in Fig. 4(c), user will enter the comparing panel in Fig. 4(d). From the comparing table, users can easily figure out the discrepancies for the corresponding vulnerability key aspects of the same vulnerability among 4 selected databases, and hence have a full view about the vulnerability characteristics.

C. Demonstration of Searching Statistics of Vulnerability Characteristics using our Website

Fig. 5(a) shows the home page of our website. Initially, the central content panel will randomly show some CVE entries. Assume we want to search statistics of vulnerability characteristics about Intel Core processors, then we can either input keyword “Intel” or “Intel Core” to the searching bar at the top and click searching button. Here we input keywords “Intel”, then the Fig. 5(b) shows the searching results. Each result card contains condensed contents including CVE ids, condensed descriptions and other information for a quick view by users.

Using multi-faceted panel on the left, we can easily filter out CVEs by vulnerability key aspects and CVSS metrics. Here we can select the product name “Core” in the product selection list, and then click “submit”. Fig. 5(c) shows the filtering results.

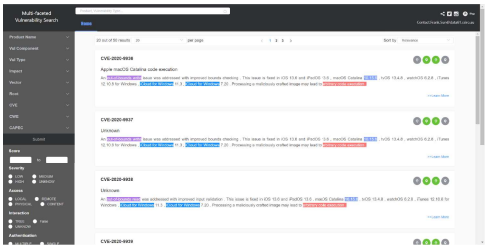
By clicking “learn more” at the bottom of CVE entry card, we can jump to content view, which is shown in Fig. 5(d). Here we show the detailed information for the CVE, including vulnerability key aspects, the sources of key aspects, and details of CVSS metrics. User can have detailed views about specific CVE.

By clicking “dashboard” at the top of content view, we can jump to dashboard page. Here we can select the statistics of specific kind of key aspects. Assume we click the product key aspects, and select target “Intel Core” in the selection menu, then we can see the statistics shown in Fig. 5(e). Here we use word clouds to present the most frequent vulnerability key aspects of the product “Intel Core”. For “Intel Core”, the most frequent vulnerable version, vulnerable component, root cause, attack vector and impact are Intel Core i3 series, information disclosure, Intel firmware, insufficient memory protection, local access, and gain elevate privileges, respectively. At the bottom, we also show the statistics of CVSS metrics of product

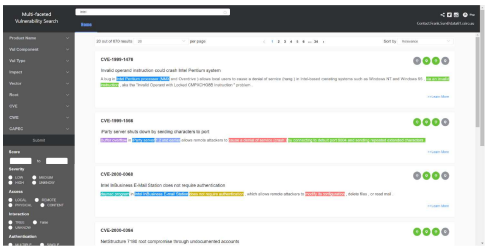
[2]https://docs.python.org/3/library/xml.etree.elementtree.html
[3]https://spacy.io/
[4]https://vuejs.org/
[5]https://www.djangoproject.com/

"Intel Core" in history, including access vector, complexity, authentication, confidentiality, integrity, availability, publish date, timeline and accumulation of CVSS scores in the history by years.

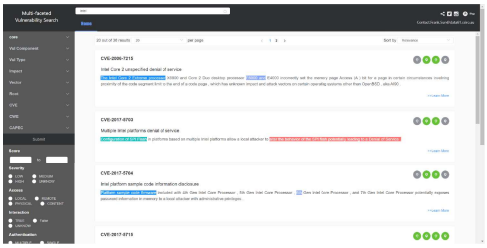
By clicking "compare" at the top, we can jump to the comparing table page, which is shown in Fig. 5(f). Here user can have a look about BERT-extracted vulnerability key aspects from 4 selected databases and their differences, so user can have a full view about discrepancies among heterogeneous databases.



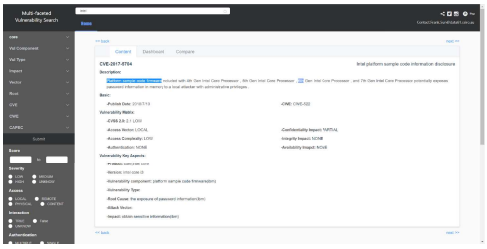
(a) Interface of Home Page



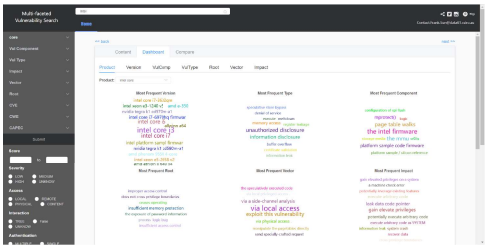
(b) Searching Results of Intel



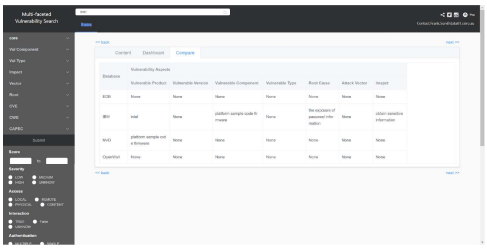
(c) Filtering Results of Intel Core



(d) Content Page of CVE-2017-5704



(e) Dashboard Page of CVE-2017-5704



(f) Compare Page of CVE-2017-5704

Fig. 5: Different Interfaces of Multi-faceted Vulnerability Searching Website