

**Table 1** p-Value of Statistical Test: homogeneous KG versus heterogeneous KG.

|               | Top-5 Acc<br>(Raw) | Top-5 Acc<br>(Filtered) | Top-10 Acc<br>(Raw) | Top-10 Acc<br>(Filtered) | MRR<br>(Raw) | MRR<br>(Filtered) | MAP      | TC-Acc   |
|---------------|--------------------|-------------------------|---------------------|--------------------------|--------------|-------------------|----------|----------|
| CWE-only KG   | 0.005859           | 0.003906                | 0.008004            | 0.001953                 | 0.064453     | 0.011407          | 0.021303 | 0.001953 |
| CAPEC-only KG | 0.009091           | 0.005889                | 0.005889            | 0.005761                 | 0.052773     | 0.040555          | 0.028954 | 0.005510 |
| CWE&CAPEC KG  | 0.009766           | 0.013672                | 0.013672            | 0.003906                 | 0.003906     | 0.001953          | 0.003906 | 0.001953 |

**Table 2** p-Value of Statistical Test: M-7 versus M-1 on overall performance of link prediction.

| Top-5 Acc<br>(Raw) | Top-5 Acc<br>(Filtered) | Top-10 Acc<br>(Raw) | Top-10 Acc<br>(Filtered) | MRR<br>(Raw) | MRR<br>(Filtered) | MAP      |
|--------------------|-------------------------|---------------------|--------------------------|--------------|-------------------|----------|
| 0.013794           | 0.001953                | 0.003906            | 0.001953                 | 0.005761     | 0.005510          | 0.005761 |

**Table 3** p-Value of Statistical Test: M-7 versus M-1 on overall performance Top-5 Accuracy (filtered) by relationship type and predict-head/tail of link prediction.

|      | instanceof | targetof | childof  | parentof | peerof   | canfollow | canprecede | semantic | overall  |
|------|------------|----------|----------|----------|----------|-----------|------------|----------|----------|
| Head | 0.001953   | 0.014369 | 0.009091 | 0.009030 | 0.034611 | 0.014147  | 0.034097   | 0.009091 | 0.005761 |
| Tail | 0.001953   | 0.009091 | 0.005666 | 0.001953 | 0.009091 | 0.014029  | 0.005857   | 0.005761 | 0.013672 |

**Table 4** p-Value of Statistical Test: M-7 versus M-1 on detailed and overall performance of triplet classification.

| instanceof | targetof | childof  | parentof | peerof   | canfollow | canprecede | semantic | overall  |
|------------|----------|----------|----------|----------|-----------|------------|----------|----------|
| 0.005793   | 0.013911 | 0.009091 | 0.005541 | 0.009091 | 0.005793  | 0.005761   | 0.007882 | 0.005793 |

**Table 5** p-Value of Statistical Test: M-7 versus M-1 on prediction performance of  $\langle CVE, instanceof, CWE \rangle$  for newly published CVEs.

| Top-5 Acc | Top-10 Acc | MRR      | MAP      | TC-Acc   |
|-----------|------------|----------|----------|----------|
| 0.001953  | 0.003906   | 0.003906 | 0.001953 | 0.005761 |