IDASH PRIVACY & SECURITY WORKSHOP 2016
-- COMPETITION RESULTS

Competition organizers:
• Haixu Tang (Indiana University)
• XiaoFeng Wang (Indiana University)
• Shuang Wang (UCSD)
• Xiaoqian Jiang (UCSD)
Track 1: Practical Protection of Genomic Data Sharing through Beacon Services (Privacy-preserving data dissemination)

Track 2: Privacy-Preserving Search of Similar Cancer Patients across Organizations (Secure collaboration)

Track 3: Testing for Genetic Diseases on Encrypted Genomes (Secure outsourcing)
Background: The Beacon project was created by the Global Alliance for Genomics and Health (GA4GH) as a means of “testing the willingness of data holders to share genetic data in the simplest technical context – query for the presence of a specified nucleotide at a given position (an allele) within a chromosome” from any human individual in a group (e.g., with a certain disease).

- >200 projects are participating the Beacon project to share their human genomic data
- Shringarpure and Bustameante recently proposed an inference attack, showing that given an individual’s whole genome sequence, an adversary may infer the presence of the individual in a beacon through repeated queries for variants in the individual's genome.

Challenge: Given a sample Beacon database, we challenge each participating team to develop a solution to mitigate the Shringarpure-Bustamante attack, while responding a maximum number of queries.

- Each team should prepare a program that responds to variation queries to any Beacon.
- The evaluation team will evaluate the submitted programs using a Beacon that was NOT shared with the participating teams.
TRACK 1: EVALUATION CRITERIA

- General criterion: the maximum number of correct queries that an algorithm can respond before any individual in the beacon can be re-identified by the Bustamante attack.

- Procedure: we perform a (modified) Shringarpure-Bustamante attack on a beacon consisting of 500 genomes extracted from the 1000 Genomes project, through the responses from each submitted program to the queries of randomly sampled variations in the Beacon.

  - We recorded the number of correct responses (and neglected incorrect responses) until the attack power reaches 0.6.

  - The error rate is computed as: \( \frac{\text{# of correct responses}}{\text{total # of queries}} \)

  - The (modified) Shringarpure-Bustamante attack utilizes allele frequencies derived from the 1000 genomes project instead of those following a presumed distribution of allele frequencies

  - Only the variations in the Beacon were queried because variations not in the database contribute little identification power for the Bustamante attack
BASELINE PERFORMANCE OF TRACK 1

- Mask k% rare SNPs the database
  - Error rate: 0.2
  - Attack power reaches 0.6 when 40,000 queries perform
  - Correctly answered queries: 32,000

- Error rate: 0.18
- Attack power reaches 0.6 when 10,000 queries perform
- Correctly answered queries: 8,200
**Background:** We consider a secure collaboration project involving two biomedical institutions: one institution hosts a sequence database of the same gene from multiple patients, and the other institution has the sequence of the gene from a single patient and wants to search it against the database to identify the patients with the top-k most similar sequences (k is typically small, <5). However, each of these two institutions cannot release their sequence data to the other institution.

- The gene is highly divergent among different human individuals (with 85%-95% sequence identity, e.g., the immune relevant genes).
- The sequence similarity is measured by the edit distance between a query sequence and sequences in the database. We assume the typical Secure Multiparty Computation (SMC) scenario: *no information should be leaked during the computation, except the final result.*

**Challenge:** Given a gene sequence database (on Party A) and a query sequence (on Party), we challenge each participating team to develop a two-party computation algorithm to identify the top-k most similar sequences in the database.

- The algorithm should consist of two programs, each executed on a computer of one party.
- The algorithm should meet the security guarantee of SMC.
- Approximation algorithms are allowed.
**TRACK 2: EVALUATION CRITERIA**

- **General criterion:** 1) security guarantee: the algorithms should not leak information other than the final results; 2) accuracy: the algorithm should report the correct top-k genes in most cases; 3) speed: the algorithm should run fast in a real-world environment, considering both computational and communication costs.

- **Procedure:** We evaluate the description of the algorithm submitted by each team; the algorithms leaking information other than the final results are disqualified. We then tested each qualified algorithms on a query gene (on one party) against a database consisting of 500 genes, in attempt to identify k=1, 3 and 5, respectively, most similar genes in the database. The ZNF717 (of ~3470 bps encoding a BRAB zinc-finger protein) gene sequences were used in the testing.

  - The submitted algorithms were executed on two virtual machines set at Indiana University and UCSD, respectively.

  - We repeated the experiment multiple times on several different databases, and recorded their running time and accuracy.

  - The algorithms are ranked according to 1) first their accuracy and 2) their running time.
Background: We consider a secure outsourcing scenario where an biomedical institution hopes to outsource the storage and computation (in this case the search of disease markers) of human genomic data on a public cloud. The genomic data will be stored in encrypted form on the cloud, and thus the search needs to be conducted by using a homomorphic encryption protocol.

Challenge: Given a single or multiple human genomes (in VCF format) and a genetic marker consisting a small number (<5) of variations, we challenge each participating team to develop a homomorphic encryption algorithm to encrypt the human genomes, and to test if any human genome carries the marker (i.e., containing all the variations).

- The algorithm should consist of two programs, one for the encryption (executed on a private computer at the biomedical institution) and one for the search (executed on the public cloud).
- The algorithm should meet the security guarantee of homomorphic encryption, no other information is leaked other than the final result.
TRACK 3: EVALUATION CRITERIA

- Hide data, query and access patterns from the cloud;
- Employ homomorphic encryption;
- 80bits security;
- 1 round query/reply;
- Maximum of 5 million variants per VCF file;
- Retrieve/reveal less than 20 variants during each search;
- Maximum of 100 client-side comparison
- Maximum of 200 VCF files (number of patients).

- Client-Server model (resembling a cloud DB);
- 10Mbps network link;

- Evaluation priority
  - Speed
  - Storage
  - Communication
- Track 1: Diyue Bu (Indiana University)

- Track 2: Lei Wang, Wenhao Wang, Diyue Bu (Indiana University)

- Track 3: Chao Jiang, Feng Chen, Shuang Wang, Le Trieu Phong, Xiaoqian Jiang (UCSD)
<table>
<thead>
<tr>
<th>Team(affiliation)</th>
<th>Member(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanderbilt University</td>
<td>Zhiyu Wan</td>
</tr>
<tr>
<td></td>
<td>Brad Malin</td>
</tr>
<tr>
<td>University of Manitoba</td>
<td></td>
</tr>
<tr>
<td>Iran University of Science and Technology</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Md Momin Al Aziz</td>
</tr>
<tr>
<td></td>
<td>Reza Ghasemi</td>
</tr>
<tr>
<td></td>
<td>Md Waliullah</td>
</tr>
<tr>
<td></td>
<td>Noman Mohammed</td>
</tr>
<tr>
<td>Team(affiliation)</td>
<td>Member(s)</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>IBM T.J. Watson Research Center and Bar-Ilan University, Israel.</td>
<td>Gilad Assharov, Shai Halevi, Yehuda Lindell, Tal Rabin</td>
</tr>
<tr>
<td>University of Manitoba and Zayed University</td>
<td>Md Momin Al Aziz, Dima Alhadidi, Noman Mohammed</td>
</tr>
<tr>
<td>University of Maryland</td>
<td>Xiao Wang, Jonathan Katz</td>
</tr>
<tr>
<td>Indiana University, Bloomington</td>
<td>Ruiyu Zhu, Yan Huang</td>
</tr>
<tr>
<td>Team (affiliation)</td>
<td>Member(s)</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Microsoft research</td>
<td>Kristin Lauter, Kim Laine, Hao Chen, Gizem Cetin, Peter Rindal, Yuhou (Susan) Xia</td>
</tr>
<tr>
<td>Communication and Distributed Systems, RWTH Aachen University, Germany</td>
<td>David Hellmanns, Martin, Henze, Jens Hiller, Ike Kunze, Sven Linden, Roman Matzutt, Jan Metzke, Marco Moscher, Jan Pennekamp, Felix Schwinger, Klaus Wehrle, Jan Henrik Ziegeldorf</td>
</tr>
<tr>
<td>Waseda University</td>
<td>Yu Ishimaki, Hayato Yamana</td>
</tr>
<tr>
<td>Seoul National University</td>
<td>Jung Hee Cheon, Miran Kim, Yongsoo Song</td>
</tr>
</tbody>
</table>
Registered Teams

- 13 countries
- 50+ teams
BEST-PERFORMING TEAMS & RESULTS

-- Result displayed is the best performance among team's submission of mitigation methods

- Team: Zhiyu Wan (Vanderbilt University)
  Brad Malin (Vanderbilt University)

- Result: No power presents even when 160,000 queries performed

- Error rate: 0.115

- Correctly answered queries: 141,600
BEST-PERFORMING TEAMS & RESULTS

-- Result displayed is the best performance among team's submission of mitigation methods

- Team: Md Momin Al Aziz (University of Manitoba)
  Reza Ghasemi (Iran University of Science and Technology)
  Md Waliullah (University of Manitoba)
  Noman Mohammed (University of Manitoba)

- Result: attack power reaches 0.6 when around 110,000 queries performed:
  - Error rate: 0.509
  - Correctly answered queries: 54,010
BASELINE PERFORMANCE OF TRACK 1

- Mask k% rare SNPs the database
  - Error rate: 0.2
  - Attack power reaches 0.6 when 40,000 queries perform
  - Correctly answered queries: 32,000

- Error rate: 0.18
  - Attack power reaches 0.6 when 10,000 queries perform
  - Correctly answered queries: 8,200
## TRACK 2: BEST-PERFORMING TEAMS

-- Evaluated by database with 500 patients records, run-time shown as average ± std through 5 runs

<table>
<thead>
<tr>
<th>Team</th>
<th>Members</th>
<th>Top 1 Run-time(s)</th>
<th>Top 1 Accuracy</th>
<th>Top 3 Run-time(s)</th>
<th>Top 3 Accuracy</th>
<th>Top 5 Run-time(s)</th>
<th>Top 5 Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM T.J. Watson Research Center and Bar-Ilan University, Israel.</td>
<td>Gilad Assharov, Shai Halevi, Yehuda Lindell, Tal Rabin</td>
<td>11.37 ±0.31</td>
<td>correct</td>
<td>11.41 ±0.17</td>
<td>2 or 3 correct</td>
<td>11.62 ±0.38</td>
<td>4 or 5 correct</td>
</tr>
<tr>
<td>University of Manitoba and Zayed University</td>
<td>Md Momin Al Aziz, Dima Alhadidi, Noman Mohammed</td>
<td>22.65 ±0.11</td>
<td>Not correct</td>
<td>22.99 ±0.15</td>
<td>2 correct</td>
<td>22.88 ±0.37</td>
<td>3 correct</td>
</tr>
<tr>
<td>University of Maryland</td>
<td>Xiao Wang, Jonathan Katz</td>
<td>12.93 ±1.26</td>
<td>correct</td>
<td>21 ±0.9</td>
<td>1 or 2 correct</td>
<td>30.4 ±2.93</td>
<td>2 or 3 correct</td>
</tr>
<tr>
<td>Indiana University, Bloomington</td>
<td>Ruiyu Zhu, Yan Huang</td>
<td>209.03 ±7.58</td>
<td>correct</td>
<td>273.14 ±7.02</td>
<td>All correct</td>
<td>337.79 ±6.18</td>
<td>4 or 5 correct</td>
</tr>
</tbody>
</table>
## TRACK 2: BEST-PERFORMING TEAMS

--- Evaluated by database with 500 patients records, run-time shown as average ± std through 5 runs

<table>
<thead>
<tr>
<th>Team</th>
<th>Top 1 Run-time(s)</th>
<th>Top 1 Accuracy</th>
<th>Top 3 Run-time</th>
<th>Top 3 Accuracy</th>
<th>Top 5 Run-time(s)</th>
<th>Top 5 Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texas A and M University (Dis-Qualified)</td>
<td>235.19 ±5.75</td>
<td>correct</td>
<td>335.50 ±14.65</td>
<td>All correct</td>
<td>525.29 ±5.08</td>
<td>All correct</td>
</tr>
<tr>
<td>University of Texas at Dallas</td>
<td>64.74</td>
<td>Not correct</td>
<td>68.72</td>
<td>Not correct</td>
<td>98</td>
<td>Not correct</td>
</tr>
<tr>
<td>Cybernetica AS</td>
<td>80.97 ±33.45</td>
<td>correct</td>
<td>67.47 ±5.39</td>
<td>1 correct</td>
<td>64.64 ±5.85</td>
<td>1 correct</td>
</tr>
<tr>
<td>RWTH Aachen University</td>
<td>95m</td>
<td>correct</td>
<td>&gt;105m</td>
<td>All correct</td>
<td>&gt;105m</td>
<td>All correct</td>
</tr>
</tbody>
</table>
- 1 query (4 variants) vs. 1 VCF file [10K records]
- 1 query (4 variants) vs. 1 VCF file [100K records]
- 1 query (1 variant) vs. 50 VCF files [100K records]
<table>
<thead>
<tr>
<th>Teams</th>
<th>The setup time [including key generation, database encryption, and upload] (s)</th>
<th>Size of the encrypted DB (MB)</th>
<th>Time to compare the query and the encrypted DB (s)</th>
<th>Memory usage of the server (MB)</th>
<th>Time to decrypt the results (s)</th>
<th>Size of the encrypted results (MB)</th>
<th>Total turnaround time [compare + transfer + decrypt] (s)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft</td>
<td>0.84</td>
<td>3.8</td>
<td>0.56</td>
<td>80</td>
<td>0.025</td>
<td>0.644</td>
<td>1.10</td>
<td>1</td>
</tr>
<tr>
<td>SNU</td>
<td>47.41</td>
<td>4.0</td>
<td>8.49</td>
<td>164</td>
<td>0.002</td>
<td>2</td>
<td>10.09</td>
<td>2</td>
</tr>
<tr>
<td>COMSYS</td>
<td>32.35</td>
<td>255.0</td>
<td>15.16</td>
<td>90</td>
<td>0.670</td>
<td><strong>0.434</strong></td>
<td>16.18</td>
<td>3</td>
</tr>
<tr>
<td>EPFL</td>
<td>137.03</td>
<td>146.8</td>
<td>6.846</td>
<td>386</td>
<td>9.366</td>
<td>3.998</td>
<td>19.41</td>
<td>4</td>
</tr>
<tr>
<td>NTU</td>
<td>619.94</td>
<td>1242.0</td>
<td>55485.6</td>
<td>1790</td>
<td>0.600</td>
<td>1.2</td>
<td>55487.20</td>
<td>7</td>
</tr>
<tr>
<td>IBM</td>
<td>538.29</td>
<td>1660.0</td>
<td>1177.59</td>
<td>3807</td>
<td>230.250</td>
<td>23</td>
<td>1426.24</td>
<td>6</td>
</tr>
<tr>
<td>WU</td>
<td>40.92</td>
<td>549.9</td>
<td>933.771</td>
<td>1448</td>
<td>0.061</td>
<td>1.558</td>
<td>935.08</td>
<td>5</td>
</tr>
<tr>
<td>Teams</td>
<td>The setup time [including key generation, database encryption, and upload] (s)</td>
<td>Size of the encrypted DB (MB)</td>
<td>Time to compare the query and the encrypted DB (s)</td>
<td>Memory usage of the server (MB)</td>
<td>Time to decrypt the results (s)</td>
<td>Size of the encrypted results (MB)</td>
<td>Total turnaround time [compare + transfer + decrypt] (s)</td>
<td>Rank</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Microsoft</td>
<td>1.86</td>
<td>24</td>
<td>3.09</td>
<td>224</td>
<td>0.024</td>
<td>0.644</td>
<td>3.6292</td>
<td>1</td>
</tr>
<tr>
<td>SNU</td>
<td>51.02</td>
<td>10</td>
<td>21.1003</td>
<td>340</td>
<td>0.00495</td>
<td>5</td>
<td>25.11</td>
<td>4</td>
</tr>
<tr>
<td>COMSYS</td>
<td>34.9</td>
<td>255</td>
<td>15.28</td>
<td>90</td>
<td>0.68</td>
<td>0.444</td>
<td>16.32</td>
<td>2</td>
</tr>
<tr>
<td>EPFL</td>
<td>137.6</td>
<td>147</td>
<td>6.79</td>
<td>3846</td>
<td>9.28</td>
<td>3.99</td>
<td>19.26</td>
<td>3</td>
</tr>
<tr>
<td>NTU</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>7</td>
</tr>
<tr>
<td>IBM</td>
<td>478.1</td>
<td>1660</td>
<td>959.1</td>
<td>3713</td>
<td>200.7</td>
<td>23</td>
<td>1178.2</td>
<td>5</td>
</tr>
<tr>
<td>WU</td>
<td>109.721</td>
<td>5447.82</td>
<td>8937.51</td>
<td>2779</td>
<td>0.05776</td>
<td>1.56</td>
<td>8938.81</td>
<td>6</td>
</tr>
<tr>
<td>Teams</td>
<td>The setup time [including key generation, database encryption, and upload] (s)</td>
<td>Size of the encrypted DB (MB)</td>
<td>Time to compare the query and the encrypted DB (s)</td>
<td>Memory usage of the server (MB)</td>
<td>Time to decrypt the results (s)</td>
<td>Size of the encrypted results (MB)</td>
<td>Total turnaround time [compare + transfer + decrypt] (s)</td>
<td>Rank</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------</td>
<td>--------------------------------</td>
<td>---------------------------------</td>
<td>------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Microsoft</td>
<td>36.69</td>
<td>188</td>
<td>32.77</td>
<td>83.6</td>
<td>1.21</td>
<td>32</td>
<td>59.58</td>
<td>1</td>
</tr>
<tr>
<td>SNU</td>
<td>2384</td>
<td>244</td>
<td>129.28</td>
<td>85.70</td>
<td>0.03218</td>
<td>122</td>
<td>226.9</td>
<td>2</td>
</tr>
<tr>
<td>COMSYS</td>
<td>1207.07</td>
<td>13000</td>
<td>278.81</td>
<td>96.264</td>
<td>0.79</td>
<td>22</td>
<td>297.2</td>
<td>4</td>
</tr>
<tr>
<td>EPFL</td>
<td>6903.1</td>
<td>1468</td>
<td>122</td>
<td>3855</td>
<td>127</td>
<td>49.97</td>
<td>288.9</td>
<td>3</td>
</tr>
<tr>
<td>NTU</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>6</td>
</tr>
<tr>
<td>IBM</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>6</td>
</tr>
<tr>
<td>WU</td>
<td>2102.28</td>
<td>27491.11</td>
<td>12447.98</td>
<td>72003.1</td>
<td>23.17</td>
<td>77.92</td>
<td>12533.48</td>
<td>5</td>
</tr>
<tr>
<td>Teams</td>
<td>1 query (4 variants) / 1 VCF [10k]</td>
<td>1 query (4 variants) / 1 VCF [100k]</td>
<td>1 query (1 variant) / 50 VCF [100k]</td>
<td>Overall score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total turnaround time (s)</td>
<td>Rank</td>
<td>Total turnaround time (s)</td>
<td>Rank</td>
<td>Total turnaround time (s)</td>
<td>Rank</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microsoft</td>
<td>1.10</td>
<td>1</td>
<td>3.6292</td>
<td>1</td>
<td>59.58</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SNU</td>
<td>10.09</td>
<td>2</td>
<td>25.11</td>
<td>4</td>
<td>226.9</td>
<td>2</td>
<td>2.67</td>
<td></td>
</tr>
<tr>
<td>COMSYS</td>
<td>16.18</td>
<td>3</td>
<td>16.18</td>
<td>2</td>
<td>297.2</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>EPFL</td>
<td>19.41</td>
<td>4</td>
<td>19.26</td>
<td>3</td>
<td>288.9</td>
<td>3</td>
<td>3.33</td>
<td></td>
</tr>
</tbody>
</table>
Special issue in BMC Medical Genomics

- Peer-review
- Submission deadline: Dec-31-2016
- Notice of decision: Jan-31-2017

http://humangenomeprivacy.org/2016/paperSubmission
- Human Longevity Inc. and GeneCloud for providing cash awards.

- NIH grants (U54HL108460, R13HG00907201A1) to support the competition

Thanks for the Participation
- Zhiyu Wan, Brad Malin, (Vanderbilt University)

- Md Momin Al Aziz (University of Manitoba), Reza Ghasemi (Iran University of Science and Technology), Md Waliullah, Noman Mohammed, (University of Manitoba)
- Gilad Asharov (Cornell), Shai Halevi (IBM), Yehuda Lindell (Bar-Ilan University), Tal Rabin (IBM)

- Md Momin Al Aziz, Dima Alhadidi*, Noman Mohammed, (University of Manitoba, *Zayed University)

- Xiao Wang, Jonathan Katz, (University of Maryland)

- Ruiyu Zhu, Yan Huang, (Indiana University, Bloomington)
- Kristin Lauter, Kim Laine, Hao Chen, (Microsoft research), Gizem Cetin, (Worcester Polytechnic Institute), Peter Rindal, (Oregon State University), Yuhou (Susan) Xia, (Princeton University)

- Jung Hee Cheon, Miran Kim, Yongsoo Song, (Seoul National University)

- David Hellmanns, Martin, Henze, Jens Hiller, Ike Kunze, Sven Linden, Roman Matzutt, Jan Metzke, Marco Moscher, Jan Pennekamp, Felix Schwinger, Klaus Wehrle, Jan Henrik Ziegeldorf, (RWTH Aachen University, Germany)

- João Sá Sousa, (EPFL) Cédric Lefebvre, (Université Toulouse), Zhicong Huang, Jean Louis Raisaro, Florian Tramer, (EPFL) Carlos Aguilar, (Université Toulouse), Jean-Pierre Hubaux, (EPFL), Marc-Olivier Killijian, (Université Toulouse)