



Visualisation of Mobile App Usage

Matthew James O'Hare - 2255357o



Project Overview

Motivation:

- 2.87 billion mobile users across the world
- Statistical methods limits the understanding that can be achieved

Aims:

- Visualise mobile app usage
- Compare dimensionality reduction algorithms



Design

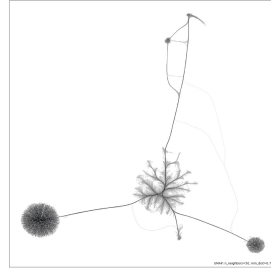
Requirements:

- Algorithms
- Pre-processing data
- Distance metric
- Clustering
- Evaluation

Technologies

Implementation

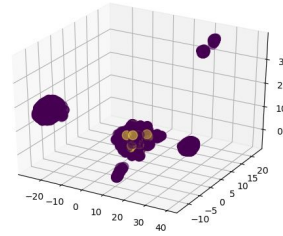
- Pre-processing data
- Distance metric
- Clustering
- Algorithms
- Command line execution



Connectivity Layout



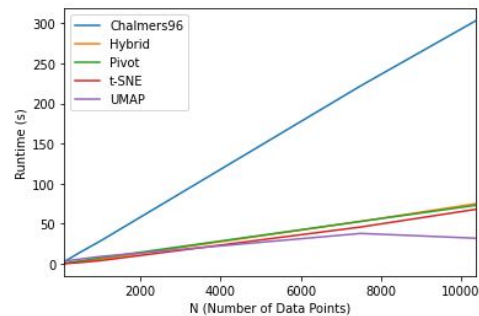
Interactive Layout



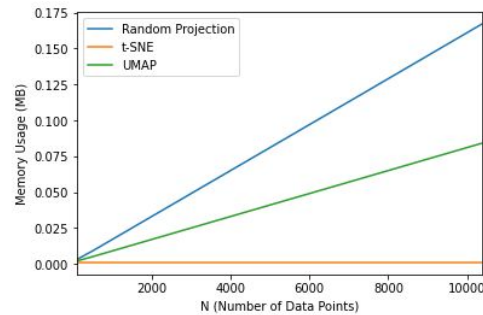
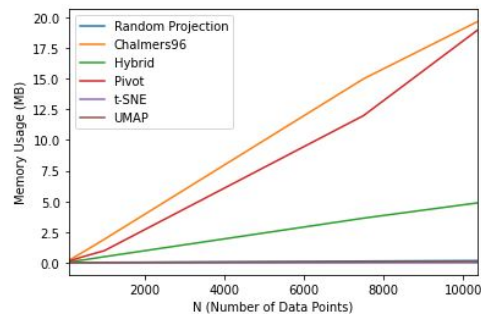
3D Layout

Evaluation

- Visual Layout
- Runtimes
- Memory Usage



Runtimes



Memory Usage

```
(base) matt@matt-pc:~/dissertation/src/algorithms$ python3 umap_layout.py
```

```
usage: python3 umap_layout.py *REQUIRED* <num apps> <dataset size> *OPTIONAL* <metric> <type> <high dimensional> <clusters>
```

```
Apps: see datasets/app_usage
```

```
Sizes: see datasets/app_usage
```

```
Metric: seucclidean, euclidean, hamming
```

```
Type: default, connectivity, diagnostic, interactive1, interactive2, 3d
```

```
High dimensional clusters: true, false
```

```
Clusters: 1 - 10
```

```
(base) matt@matt-pc:~/dissertation/src/algorithms$ python3 umap_layout.py 100 7504
```

```
Creating default layout of 7504 app usage entries using a metric of seucclidean with 2 clusters. High dimensional clusters - True
```

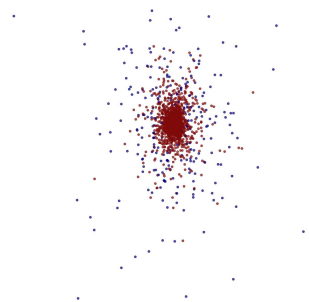
```
Runtime: 39.291998518569946
```

```
Layout time: 39.46s (0.7 mins)
```

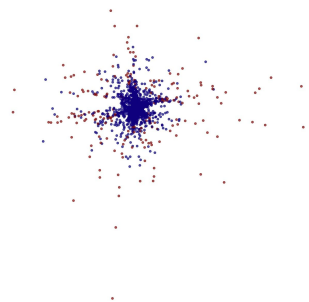
```
(base) matt@matt-pc:~/dissertation/src/algorithms$ _
```

I

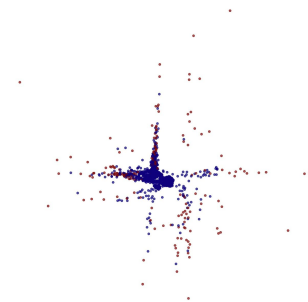
Other Layouts



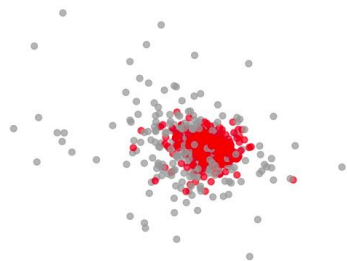
Chalmers' 1996



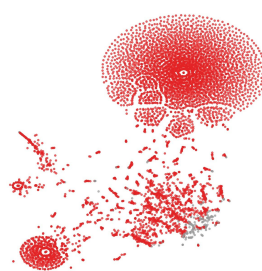
Hybrid Layout



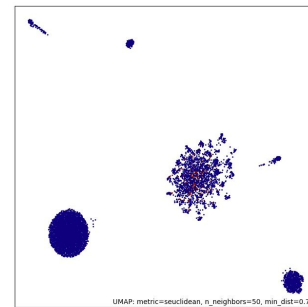
Pivot Layout



Random Projection



t-SNE



UMAP



Conclusion

- What this project achieved
- Key point: never underestimate the importance and effectiveness of a carefully planned and thought out process