Andrew Twigg

Belmont, CA 94002
andy@atwigg.com

I'm a computer scientist with expertise in algorithms, databases and machine learning. I did my PhD in graph algorithms at Cambridge (2006), was at the CS dept at Oxford 2008-13 and have been founder/CTO of 2 acquired startups.

Employment

2020-present Distinguished Engineer, Lacework
I work on ML and large-scale data processing efforts at Lacework. We collect billions of messages per hour from agents and cloud data sources, and use a variety of techniques including unsupervised graph representation learning and anomaly detection.

2018-2020 Entrepreneur in residence, Milliways/Rocketship.vc
EIR role, exploring ideas around deep RL. One project was to train a high-frequency trading agent to learn how to exploit market microstructure. We collected several TBs of level 3 messages from a large cryptocurrency exchange. This was able to do full nanosecond LOB reconstruction and simulation for better queue length estimation, etc. We helped Stanford students explore similar ideas via CS246 coursework.
https://medium.com/@andytwigg/learning-to-trade-with-deep-rl-666ed6bbd921

2015-18 Chief Scientist, Insidesales.com
2014-15 CTO, C9 (acquired by Insidesales.com)
We were one of the first companies to apply ML to CRM sales data to to improve efficiency of sales teams. C9 was acquired by Insidesales.com which grew to over 800 people and raised over $250m in funding.
https://blogs.gartner.com/todd-berkowitz/insidesales-com-buys-c9-predictive-sales-analytics/
https://blogs.wsj.com/venturecapital/2015/05/28/insidesales-com-acquires-c9-in-predictive-analytics-battle/

2013 Founder, Featurestream.io
I built a novel streaming, distributed random forest implementation on spark streaming and experimented with offering it as a hosted API service.
https://medium.com/@andytwigg/featurestream-io-random-forests-6992b03b521
https://github.com/featurestream/featurestream-client

2009-12 CTO & cofounder, Acunu (acquired)
Acunu built a streaming data analytics engine on Cassandra based on novel approaches to write-optimized materialized views ("stratified B-trees") with SSD-friendly implementations with low write amplification. We started the London Big Data meetup, filed several patents and raised over $10m in funding. Company was acquired in 2013.
https://arxiv.org/abs/1103.4282
Persistent Cache-oblivious Streaming Indexes

2008-13 Oxford CS Department & Fellow, St. John's College, Oxford

Elected JRF by open competition. Research interests: efficient algorithms for data mining via approximation, streaming, machine learning. Took a sabattical to found Acunu.

2006-7 Microsoft Research, Cambridge and Technicolor Research, Paris
Worked with Laurent Massoulie on randomized peer-to-peer algorithms for broadcasting and live streaming problems with provably optimal throughput.
Randomized decentralized broadcasting algorithms (INFOCOM 2007)

Education

2006 PhD Computer Science, Cambridge University (King's College)
PhD Thesis: Forbidden-Set Compact Routing (advisor: Ken Moody)
I worked on graph algorithms for "compact routing" problems. The forbidden-set distance problem deals with how to precompute small labels in a graph that allows one to approximate distances in the presence of unknown failures. We used this to develop the first low-stretch failure-tolerant compact routing protocol. Nominated for British Computer Society Best Dissertation Award.

1999-2002 BSc Computer Science, Warwick University
Thesis: Lightweight web prediction algorithms (advisor: Prof Mike Paterson)
Awarded top 1st result, best overall student each year and top overall graduating student

Teaching

2008-11 Oxford University, Computer Science/Mathematics
Randomized Algorithms (Masters), Advanced Data Structures and Algorithms

2003-6 Cambridge University, Computer Science
Algorithms and Data Structures, Probability, Complexity Theory

Interests

I enjoy playing drums (grade 8), golf and I rowed for Cambridge CULRC and King's 1st VIII.

Selected Publications

Persistent Cache-oblivious Streaming Indexes. ArXiv, abs/1707.08186, 2017

Locality-preserving allocations problems and coloured bin packing. Twigg, A and Xavier, E. Theoretical Computer Science, 596:12-22, 2015

Stratified B-trees and versioned dictionaries. Twigg et al, HotStorage'11, Berkeley CA

Optimal query/update tradeoffs in versioned dictionaries. Twigg, A, Byde, A ArXiv, abs/1103.2566, 2011 *https://arxiv.org/abs/1103.2566*

Constrained-path labellings on graphs of bounded clique-width. Twigg,A, Courcelle,B Theory Comput. Syst., 47(2):531–567, 2010

Epidemic live streaming: optimal performance trade-offs. Bonald et al. SIGMETRICS,2008

Worst-case time decremental connectivity and k-edge witness problems. ArXiv,abs/0810.5477, 2008

Connectivity checking in 3-connected planar graphs with obstacles. Courcelle et al. Electronic Notes in Discrete Mathematics, 31:151–155, 2008

Rate-optimal schemes for peer-to-peer live streaming. With Laurent Massoulie. Journal Performance Evaluation, 65(11-12):804–822, 2008

Randomized decentralized broadcasting algorithms. Laurent Massoulie, Andrew Twigg, Christos Gkantsidis, and Pablo Rodriguez. In INFOCOM, pages 1073–1081, 2007

Forbidden-set labelling on graphs. With Courcelle et al. PODC workshop on Locality Preserving Distributed Computing Methods (LOCALITY), 2007

Compact forbidden-set routing. Bruno Courcelle and Andrew Twigg. STACS 2007.

The complexity of fixed point models of trust in distributed networks. Karl Krukow and Andrew Twigg. Theoretical Computer Science, 389(3):528–549, 2007

Compact forbidden-set routing (PhD Thesis). Technical report UCAM- CL-TR-678, 2006

Provably optimal decentralized broadcasting algorithms. With Massoulie et al, Technical report, 2006. MSR-TR- 2006-105