FIMS: Identifying, Predicting and Visualising Food Insecurity

Georgiana Nica-Avram
psxga@nottongham.ac.uk
University of Nottingham
Nottingham, United Kingdom

John Harvey
john.harvey2@nottongham.ac.uk
University of Nottingham
Nottingham, United Kingdom

James Goulding
james.goulding@nottongham.ac.uk
University of Nottingham
Nottingham, United Kingdom

Benjamin Lucas
benjamin.lucas@nottongham.ac.uk
University of Nottingham
Nottingham, United Kingdom

Andrew Smith
andrew.p.smith@nottongham.ac.uk
University of Nottingham
Nottingham, United Kingdom

Gavin Smith
gavin.smith@nottongham.ac.uk
University of Nottingham
Nottingham, United Kingdom

Bertrand Perrat
bertrand.perrat@nottongham.ac.uk
University of Nottingham
Nottingham, United Kingdom

ABSTRACT

Food insecurity is a persistent and pernicious problem in the UK. Due to logistical challenges, national food insecurity statistics are unmeasured by government bodies - and this lack of data leads to any local estimates that do exist being routinely questioned by policymakers. We demonstrate a data-driven approach to address this issue, deriving national estimates of food insecurity via combination of supervised machine learning with network analysis of user behaviour, extracted from the world's most popular peer-to-peer food sharing application (OLIO). Despite long-standing theoretical links between social graph topologies and physical neighbourhoods, prior research has not considered dimensions of geography, network interactions and behaviours in the digital/analogue space simultaneously. In addressing this oversight, we produce a browser-based, interactive and rapidly updateable visualisation, which can be used to analyse the spatial distribution of food insecurity across the UK, and provide new perspective for policy research.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning approaches**; • **Applied computing** → **Sociology**; • **Human-centered computing** → **Geographic visualization**.

KEYWORDS

Food Insecurity, Hunger, Poverty, Geospatial, Data, Visualization, Computational Social Science

1 INTRODUCTION

Food insecurity is a remarkably persistent problem in the United Kingdom. Despite being the fifth wealthiest country in the world [2], inequality endures to such an extent that many people cannot afford basic provisioning. Recent research [1, 4, 9] suggests an increasing number of people are facing food hardship, many experiencing 'in-work poverty' where salaries fail to cover even basic expenditure. Though the UK government measures an *Index of Multiple Deprivation* (IMD), which includes income, employment, and health, there exists a dearth of reliable large-scale data on the number of people unable to reliably procure enough food for themselves and their families [3].

Food insecurity surveys, often deployed in other countries, are not currently used in the UK. Such surveys are time-consuming, expensive, and practically challenging to implement given the nature of their focus. The resulting lack of data, however, not only hinders progress on public policy in the UK, but renders local statistics which do exist susceptible to questioning (such as those available from organisations such as food banks, who routinely report growing numbers year on year) - allowing leading politicians to deny their validity and/or argue that they are either unreliable or politically motivated, e.g. [14].

The UK is not unique in lacking publicly available sources of longitudinal data to measure these issues. Worldwide, food waste and food insecurity are persistent problems at the core of two sustainable development goals and alleviating them should be treated as a 'global joint political effort' for all stakeholders, from governments, civil society, to businesses and individual consumers [6, p.1]. However, efforts often fall short because of a lack of shared understanding of how these Sustainable Development Goals can be operationalised [12]. A lack of knowledge 'on feedback between social and ecological systems' can inhibit monitoring and intervention [10, p.1115]

In the UK, the lack of reliable food insecurity estimates need not be the case, especially given the growing popularity of, and data

embedded in, peer-to-peer food sharing networks. We demonstrate a methodological approach to deriving national estimates of food insecurity via combination of supervised machine learning with network analysis of user behaviour, extracted from the world's most popular peer-to-peer food sharing application (OLIO). Despite long-standing theoretical links between neighbourhoods and food sharing social networks, the simultaneous linking of physical and digital dimensions has not previously been undertaken [13]. Data from OLIO's peer-to-peer food sharing platform can help bridge the gap between neighbourhood characteristics, such as economic deprivation or access to food stores, with geography and evidence provided from social interactions.

The Food Insecurity Modelling system (FIMS) we present seeks to remedy this important issue of obtaining rapidly updateable food poverty estimates from behavioural data. This approach represents an alternative in the absence of government led measurement. The system integrates dynamics of social networks and local geographies, to produce food insecurity estimates at fine grained neighbourhood levels.

The following sections outline (1) the method used to classify and subsequently predict food insecurity within FIMS, applying supervised machine learning on proprietary data drawn from the OLIO food-sharing application; (2) how observed and inferred data around food insecurity can be aggregated and visualised at a national level; and (3) the potential contribution of this new approach and a discussion of future research opportunities to help tackle the challenge of measuring food insecurity to inform policy. The full Food Insecurity Modelling system (FIMS) we present is illustrated in Figure 1.

2 IDENTIFYING AND PREDICTING FOOD INSECURITY

In the absence of government-led primary data collection, there are limited options for measuring food insecurity across the UK. However, as food supply chains become increasingly digitised, new data sources are starting to emerge. This is especially true where people use digital technology to mediate their consumption patterns, for example peer-to-peer food surplus sharing. Not only do these networks encode important behaviours, such as the sharing, or soliciting of food, but they also capture many real-world interactions. Live data streams open the possibility to study how consumer behaviour changes over time rather than sporadic snapshots gained through surveys. They also open the possibility to model and visualise individuals' standing in a broader network of neighbours, and the co-occurrence of socioeconomic factors. Mapping social and environmental processes, in particular, helps reduce the complexity around food insecurity [7].

OLIO is the world's leading peer-to-peer food sharing application. Over 1.5 million users have signed up to the service across 49 countries. It is important to state that OLIO is not designed to solve poverty or food insecurity, the aim of the application is to reduce food waste of consumers and businesses. However, recent research has shown, that as with the UK population in general, there is a sub-population experiencing food insecurity. Drawing from OLIO's datasets, which simply logs the food listings, requests and activity that users make publicly available through the app, FIMS is able
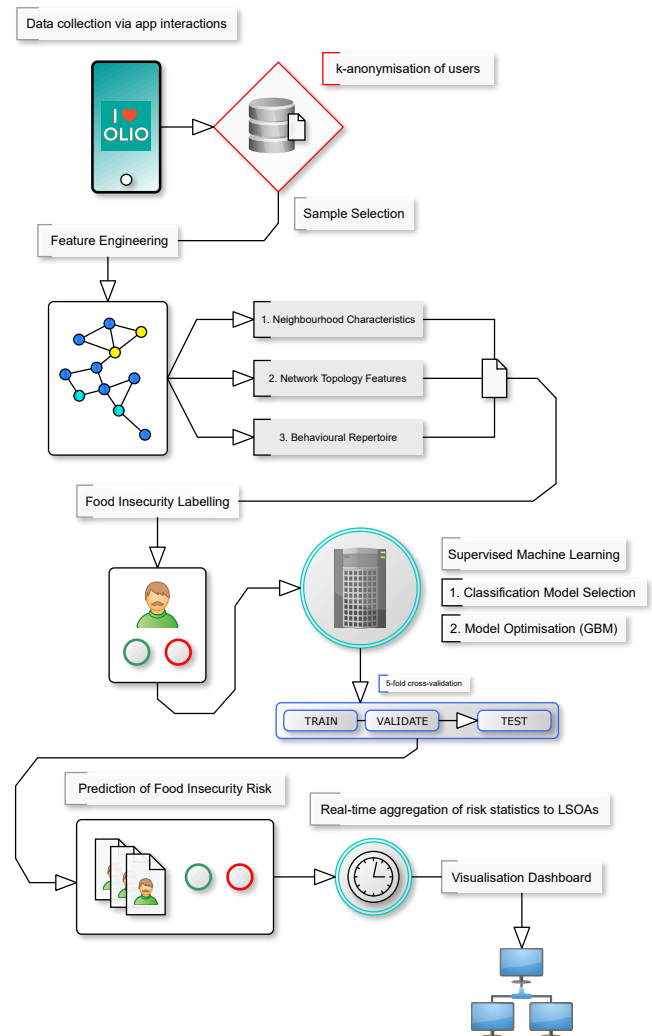

Figure 1: Overview of FIMS system workflow, resulting in a visualisation of food insecurity estimates across the UK

to model and predict the prevalence of food insecurity following three key steps:

- Pseudonymized user modelling across three core dimensions: (1) neighbourhood deprivation characteristics, (2) food-sharing metrics reflected via OLIO, and (3) network features, reflecting location within the food sharing network topology as a whole.
- Via a labelled dataset of individuals self-declared as in food hardship, stratified modelling of a generalised classifier is used to identify instances of food insecurity.
- The resulting model can then be applied to the whole sample population to identify the prevalence of food insecurity vulnerability. These statistics can then be aggregated into geographic clusters, resulting in food-insecurity predictions from local to national levels.
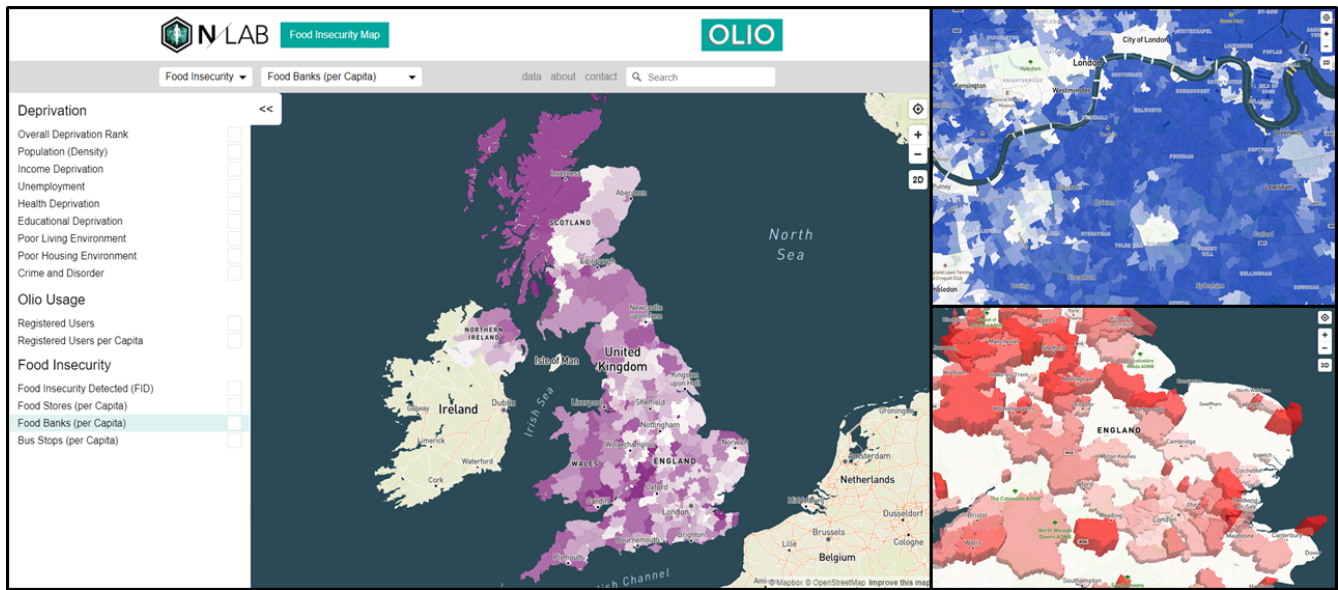
Figure 2: Left: Overview of food insecurity map interface - showing UK results at Local Authority District level. Top right: IMD statistics visualised at LSOA level. Bottom Right: Relative food insecurity predictions, visualised at MSOA level and rendered in 3D to aid comparison

2.1 Feature Engineering

Feature engineering involved derivation of 35 variables, derived over three user dimensions. 15 were derived directly from activity data (from 'listing frequencies' to 'likes'), 7 features from network analysis statistics, and 13 from socio-demographic data associated with neighbourhood location of activities). This resulted in a training dataset of 52,881 user data points, after filtering to ensure at least one successful sharing or soliciting of food on OLIO, and registration to one of the UK's 42,619 neighbourhoods, technically referred to as Lower Super Output Areas (LSOAs).

To ensure privacy, the socio-demographic features associated with each data-point were relatively coarse, and assigned from a neighbourhood level, using census measures of deprivation (n.b. each LSOA averaging 1,200 residents). Measurements of access to around 12,000 food stores, 2,212 food banks and centres and more than 400,000 bus stops were also calculated for each LSOA, and attributed to corresponding data-points. For food-sharing activity characteristics the system summarised users' food sharing levels, as well as how much food they solicited and how bursty or sporadic their interactions were. Finally, for the network topology dimension a descriptive social network analysis (via Python's NetworkX library) was used to derive measures of degree distribution, centrality and clustering. While the first two dimensions offer insights on geographic propinquity, socioeconomic affinities or personal preferences, social network analysis reveal further insights about kinship and inter-dependencies arising in these local networks of food aid [5, 11].

2.2 Classifier Development

Classifier selection was undertaken via a supervised machine learning approach. This was possible via a labelled sub-sample of 361

users self-declaring to be experiencing food-insecurity in their food soliciting messages on OLIO. These were combines in a stratified fashion, with an equivalent sub-sample of food-secure individuals, to form our dataset for model optimisation. Four competing model classes were investigated (Support Vector Machines, Random Forests, Gradient Boosted Machines and k-Nearest-neighbour), with the performance of each model being assessed using a rigorous five-fold cross-validation regime (as per Figure 1. Based on this approach, the Adaboost model showed highest levels of precision and recall, and classification accuracies of 71%. At a 70% prediction probability threshold, 5% of users in the sample were predicted to be food-insecure, 55% - food-secure. Moreover, f1 scores for this model illustrate that these prediction levels were not the result of one class being disproportionately favoured over the other. While marginally so, f1 scores for users experiencing hardship were lower - 0.62 compared to 0.76. This resonates with insights from initial exploration of the sample, whereby users' food security changes during their time on the platform, and so does their sharing behaviour. Some start by soliciting food for themselves, then transition to acting as volunteers for OLIO, collecting food from associated stores and distributing it in the community. Finally, an extensive variable selection analysis was performed, along with summarisation via principal component analysis (PCA) and ANOVA F-value feature ranking, in order to evidence credibility of the selected model[1].

---

[1] further analysis of methods for modelling and predicting instances of food insecurity can be found in Nica-Avram *et al.* (Forthcoming)

2.3 Modelling the geographic prevalence of food insecurity

Once a final model was obtained, probabilistic classifications could be obtained for all datapoints in our sample, and estimated prevalence summarised at LSOA level (thus food insecurity data output by FIMS is both psuedo-anonymised and then aggregated at scale to remove the possibility of triangulation). We worked with the mapping company, *MapBox*, to produce an interactive web-interface, allowing visualisation of FIMS' modelling results, along with several accompanying data layers for comparison. Though the *Index of Multiple Deprivation* has been visualised previously, to our knowledge this is the first time food insecurity statistics have been estimated and visualised in the UK. OLIO is widely downloaded and used across the UK, and continues to grow, providing scope for continued updates at a national level.

The interactive tool demonstrated has been specifically designed to enable searching and comparison between different geographic resolutions including Lower Super Output Area (LSOA), Middle Super Output Area (MSOA) and Local Authority Districts. Observed and inferred data can be combined for each region to give further insight into the relative experience of deprivation and food insecurity. For example, though areas may experience comparable deprivation, they may differ in infrastructure for assisting people to find healthy, affordable food.

Food insecurity has one cause: poverty. However, the experience of food insecurity differs for people depending on the availability of nearby affordable food and emergency food assistance. Reflecting some of these accompanying measures, FIMS also provides aggregate summaries for the number food stores, food banks, and bus stops in each region. These data help to provide further contextualisation to issues that entrench food insecurity.

3 CONTRIBUTION AND FUTURE RESEARCH DIRECTIONS

This demonstration illustrates the potential for a novel, data-driven means of estimating food insecurity prevalence at a national level. Though the observed data originates from a single platform, and is necessarily biased to those with the ability to access it, the extensive user-base means that FIMS reflects much food for thought. While much work remains to be done verifying the system, FIMS produces a picture of food insecurity at a more-fine grained level than previously obtainable - and at a far greater geographical extent than can be achieved from traditional surveys (which are often highly localised). Furthermore, as the underlying model can be automated across contexts, the method is much more cost and time effective than surveys, which have historically meant research is only conducted sporadically.

Additionally, the digital/analogue interactions encoded in this data set open possibilities for linking further heterogeneous data sources. Future work should bolster the insights available by drawing upon multiple sources of anonymised aggregated data, from food consumption figures to news, for long-lasting social impact. We also suggest further validation of the approach could be gained by examining localised food insecurity survey results (e.g. recently commissioned by the Mayor of London, see [8]) in relation to the

predictions generated. Further research should also focus on questions of access, for example, what is the relation between food deserts (areas with poor access to affordable food) and food insecurity - how do they intersect? The experience of food insecurity is mediated by where people live, and as emergency food assistance and supermarkets tend to cluster around population centres, there may be notable differences in urban and rural areas. Solutions, too, require complementary actions based on improved knowledge of social and ecological systems.

REFERENCES

[1] John R. Ashton, John Middleton, and Tim Lang. 2014. Open Letter to Prime Minister David Cameron on Food Poverty in the UK. *The Lancet* 383, 9929 (2014), 1631.

[2] Credit Suisse. 2019. *Global Wealth Report 2019*. Retrieved January 2, 2020 from https://tinyurl.com/y3tr4pwt

[3] Department for Communities and Local Government. 2015. *English Indices of Deprivation 2015*. Retrieved 2018-12-01 from https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

[4] E. Dowler and H. Lambie-Mumford. 2015. Food for Thought: An Ethnographic Study of Negotiating Ill Health and Food Insecurity in a UK Foodbanks. *Social Policy and Medicine* 132 (2015), 38–44.

[5] César Ducruet and Laurent Beauguitte. 2014. Spatial Science and Network Science: Review and Outcomes of a Complex Relationship. *Networks and Spatial Economics* 14, 3-4 (2014), 297–316.

[6] Francesca Galli, Alessio Cavicchi, and Gianluca Brunori. 2019. Food Waste Reduction and Food Poverty Alleviation: A System Dynamics Conceptual Model. *Agriculture and Human Values* (2019), 1–12.

[7] Menno Kraak, Britta Ricker, and Yuri Engelhardt. 2018. Challenges of Mapping Sustainable Development Goals Indicators Data. *ISPRS International Journal of Geo-Information* 7, 12 (2018), 482.

[8] London Datastore. 2019. *The First Food Security Estimates for London*. Retrieved January 2, 2020 from https://tinyurl.com/ye7wsbop

[9] R. Loopstra, A. Reeves, D Taylor-Robinson, B. Barr, M. McKee, and D. Stuckler. 2015. Austerity, Sanctions, and the Rise of Food Banks in the UK. *British Medical Journal* 350 (2015), p.h1775.

[10] Matías E Mastrángelo, Natalia Pérez-Harguindeguy, Lucas Enrico, Elena Bennett, Sandra Lavorel, Graeme S. Cumming, Dilini Abeygunawardane, Leonardo D. Amarilla, Benjamin Burkhard, Benis N. Egoh, et al. 2019. Key Knowledge Gaps to Achieve Global Sustainability Goals. *Nature Sustainability* 2, 12 (2019), 1115–1121.

[11] Miller McPherson, Lynn Smith-Lovin, and James M. Cook. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology* 27, 1 (2001), 415–444.

[12] Jeffrey D. Sachs, Guido Schmidt-Traub, Mariana Mazzucato, Dirk Messner, Nebojsa Nakicenovic, and Johan Rockström. 2019. Six Transformations to Achieve the Sustainable Development Goals. *Nature Sustainability* 2, 9 (2019), 805–814.

[13] Martin Saveski, Eric Chu, Soroush Vosoughi, and Deb Roy. 2016. Human Atlas: A Tool for Mapping Social Networks. In *WWW'16 Companion, April 11- 15, 2016*. ACM, Montréal, Québec, Canada. https://doi.org/10.1145/2872518.2890552

[14] Peter Walker. 2017. Dominic Raab Accused of 'Stupid and Offensive' Food Bank Comments. *The Guardian* (2017). Retrieved 2019-03-06 from https://tinyurl.com/y9kwj9y5