Modeling, Analysis, and Design of Residential Energy Saving Strategies: A Systematic Investigation of Individual and Social Network Influences

PhD Thesis

Feng Du

Wind Energy and Control Centre
Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow

February 21, 2024

Abstract

Reducing energy consumption in the residential use sector could significantly improve energy savings, especially in this post-COVID era when working from home becomes a favored option for many people. Each home's energy consumption pattern depends on individual user factors and is also influenced by the networks.

Understanding the impacts of individual and social networks on household energy saving behaviours is an unexplored and challenging task. There's a lack of systematic approach to enable modelling, quantitative analysis and optimization design. This motivates the research in this PhD project. The key innovative contributions are summarized in the following.

Firstly, a small-scale network with detailed nodes connection information is studied. A network model is established, considering information diffusion within householders' social network, from which the expected energy savings (EES) can be calculated. The connection level between users are weighted by coefficients. The influence from the source node to the target node is estimated by the use of probability theory. Considering information decay with different degree separations, the propagation coefficient that measures the diffusion of EES information can be calculated dynamically, based on which nodes that produce the largest EES are identified. From this study (in Chapter 4), the role of information diffusion within a social network in promoting energy efficiency products (EEP) is found to be significant.

Secondly, factors that may affect user's EEP adoption decision are analyzed. Four factors are selected: the personal acceptance level, the influence from the connected neighbors, the overall network adoption rate, and the advertisement influence from a wider environment. Among them, the personal acceptance is a grouped factor influenced by individual factors such as household income, age group, family status, and employment status. A network model is established to integrate multiple input factors towards their impacts on EEP adoption decision. Unknown model parameters are estimated using survey data. An adoption utility measure is defined based on the four impact factors with associated weightings, through which each user's EEP adoption decision can be determined, and consequently the adoption rate of the whole network can be calculated. From this study (in Chapter 5), it is revealed that advertisement input plays a key role to influence user's EEP adoption through a social network.

Thirdly, the impact of advertisement on residential home EEP promotion is studied, with the aim to investigate the influence of advertisement control on energy savings through social networks. A mathematical model is established to predict the EES in a network where advertisement is used to influence the users' EEP adoption decision. The proposed model consists of information diffusion, EES calculation, and advertisement control. It can be applied to mass roll-out programm to forecast the EES and the adoption rate of new energy products, based on which the advertising investment required to accelerate energy savings can be determined by optimization design. Epidemic theory is employed in modeling to characterize the dynamics of information diffusion on EEP. The change of response rate and adoption rate due to the influence from advertisement is quantified using Bayesian forecasting theory (BFT). Case studies for different scenarios are conducted, considering various optimization targets such as adoption rate, time cost, advertisement cost, and total energy savings subject to program budget and time constraints. This study (in Chapter 6) indicates the potential of using advertisement as a means to promote EEP.

Two population networks are established using survey data, starting from a small network including 40 homes/nodes, followed by a large one with one million nodes. The small network is used for all three chapters (4-6), the large network is used for Chapters 5 and 6.

The integrated work of network description, information diffusion modelling, parameter identification, sensitivity analysis and optimization design provides a useful tool to analyze and manage individual and social impacts on household energy savings.

Contents

Abstract **ii**

List of Figures **viii**

List of Tables **xii**

Nomenclature **xiv**

Acronyms **xviii**

Preface/Acknowledgements **xxii**

1 Introduction **2**

1.1 Research Background 5

1.2 Problem Description 6

1.3 Research Aims and Objectives 6

1.4 Research Contribution 7

1.5 Thesis Overview 8

1.6 Publications 11

2 User Behaviour, Social Networks and Energy Savings: A Review **12**

2.1 Introduction 12

2.2 User Behaviour and Dwelling Factors on Energy Savings 14

2.2.1 Human Behaviour on Energy Savings 14

2.2.2 Dwelling Factors 16

2.3 Strategic Energy Plan 17

2.3.1 Definition and Significance 17
2.3.2 Strategic Identification 19
2.3.3 Stakeholder Engagement 19
2.3.4 Social Aspects of Energy Savings 20
2.4 User Influence on HVAC Energy Consumption 22
2.5 Occupant Behaviour and Clustering Analysis 24
2.6 Social Networks on Consumer Behaviour 26
2.6.1 Motivational Frameworks within Social Networks 27
2.6.2 Utilizing Efficient Interactions in Energy Technology Adoption . 27
2.6.3 Entropy-based Analysis in Energy Consumption 29
2.7 Summary 30
2.7.1 Research Gap 31

3 Modelling Methods for Social Networks: Review and Preliminaries 32
3.1 Introduction 32
3.2 Regular Networks 34
3.3 Random Networks 36
3.3.1 Key Features and Representative Use 36
3.3.2 Implications in Human Interaction and Information Propagation 38
3.4 Complex Networks 39
3.4.1 Small-World Networks 39
3.4.1.1 Watts-Strogatz Model: A Small-World Network Model 40
3.4.1.2 Applications of Small-World Networks 42
3.4.2 Scale-Free Networks 42
3.4.3 Comparison of Random and Complex Networks 44
3.5 Spreading Dynamics of Complex Networks 45
3.6 Bayesian Forecasting Theory (BFT) 46
3.7 Summary 47

4 Modelling and Analysis of Expected Energy Savings in Social Networks 50

4.1 Introduction 50
4.2 Network Model Development 53
4.2.1 Social Networks with Connection Weightings 53
4.2.2 Evolution of Connection Probabilities 54
4.2.3 Information Diffusion 55
4.2.4 Calculation of Installation Number 58
4.2.5 Calculation of Expected Energy Savings 59
4.2.6 Estimation of Model Parameters 60
4.3 Information Sources Selection Design 63
4.4 Case Study of a Small Scale Social Network 63
4.4.1 Small Network Survey and System Setup 63
4.4.2 Model Parameter Identification 66
4.4.3 Single Information Source 67
4.4.3.1 Evolution of Installation Number 68
4.4.3.2 EES Profiles Over Different Window Lengths 69
4.4.4 Sensitivity Analysis with Monte Carlo Calculation 75
4.4.5 Multiple Information Sources 78
4.4.6 Cost Benefit Analysis 81
4.5 Summary 83

5 Investigating Impact of Personal and Social Network Factors to Adoption of Energy Products 85

5.1 Introduction 85
5.2 Methodology on Model Development 86
5.2.1 Network Modelling 86
5.2.2 Personal Acceptance Level Model 88
5.2.3 Neighbouring Network Influence Model 89
5.2.4 Network Trend Model - Network Adoption Rate 90
5.2.5 Advertisement Influence Model 90
5.2.6 Adoption Utility 92
5.2.7 Decision of Adoption - Update of Node State 93

5.2.8 Determination of Unknown Model Parameters 95

5.3 Case Study with Two Networks 96

5.3.1 System Setup 96

5.3.2 Personal Acceptance Level 97

5.3.3 Advertisement Input Setup 100

5.3.4 Small Population Network 101

5.3.4.1 Baseline System Simulation 101

5.3.4.2 Advertisement Control Practice 104

5.3.5 Large Population Network 108

5.3.6 Parameter Sensitivity Analysis 114

5.4 Summary 118

6 Influence of Advertisement Control to Energy Savings 120

6.1 Introduction 120

6.2 Advertisement Control Model for Adoption Rate and Energy Savings . . 122

6.2.1 Epidemic Model 123

6.2.2 Information Diffusion Model without Advertisement Input 124

6.2.3 Information Diffusion Model with Advertisement Input 126

6.2.4 Update of Time-dependent Parameters in Advertisement Model . 127

6.2.5 Estimation of Model Parameters 128

6.2.6 Electricity Cost Saving Model 128

6.2.7 Advertisement Cost and Free Trial Product Cost 129

6.3 Advertisement Control Design 130

6.4 Case Studies 132

6.4.1 Experimental Settings and Parameter Estimation 133

6.4.2 Parameter Sensitivity Analysis 136

6.4.3 Advertisement Control Design for Small Population Network . . 138

6.4.3.1 Baseline Strategy 138

6.4.3.2 Method 1: Targeting for Lowest Advertisement Cost . . 139

6.4.3.3 Method 2: Targeting for Shortest Time 143

6.4.3.4 Method 3: Targeting For Largest Energy Savings 144

6.4.3.5 Method 4: Targeting for Largest Installation Number . 147
6.4.3.6 Cost Effectiveness Analysis 148
6.4.3.7 Summary for Small Network Simulation 150
6.4.4 Advertisement Control of Large Population Network 150
6.4.4.1 Baseline Strategy 150
6.4.4.2 Method 1: Targeting for Lowest Advertisement Cost . . 151
6.4.4.3 Method 2: Targeting for Shortest Time 155
6.4.4.4 Method 3: Maximize for Total Energy Savings 155
6.4.4.5 Method 4: Maximize the Adoption Rate 158
6.4.4.6 Cost Effectiveness Analysis 159
6.5 Summary 160

7 Conclusions and Future Work 163
7.1 Conclusions 163
7.2 Future Work 167
7.2.1 Potential Use of ML 168

A Large Population Network Campaign Data (Population 1 Million) 169

B Small Population Survey Questionnaire 172

Bibliography 176

List of Figures

1.1 UK Energy Consumptions by Sector 2020 (Parker, 2021) 3
1.2 Energy Consumption of UK 2020 (Parker, 2021) 4
1.3 Conceptual Thesis Flow 9

2.1 Taxonomy of User/Social Influence to Residential Home Energy 13

3.1 Social Network Models by Randomness and Heterogeneity 33
3.2 Network graphs with different randomness, from low rewire probability ($P_r = 0$) to high rewire probability ($P_r = 1$), connections before the rewiring process are in blue, after the rewired connections are in green. . 35

4.1 Energy Efficiency Information Spreading through Individual Interactions within a Social Network 51
4.2 Illustration of Information Diffusion Mechanism in a Network 56
4.3 Graph Representation of the Small-Scale Social Network 64
4.4 Convergence of a_{ij} Coefficients during the Estimation Process 66
4.5 Convergence of μ_j Coefficients during the Estimation Process 67
4.6 Number of Installations when Different Nodes are Selected as the Single Information Source 68
4.7 Sensitivity Analysis of EES Model Parameters under Four Standard Deviation Levels 76
4.8 Difference of Installation Numbers Achieved by Nominal Value and Monte-Carlo Calculation under Parameter Variation(Parameter μ_{38}, Standard Deviations: 10%, 20%, 50% and 100% of Mean Value) 77

5.1 Model Configuration 87

5.2 Decision of Adoption Algorithm Flow Chart 94

5.3 Convergence Graph along Iterations in Fmincon Function of Parameter Estimation (m_1, m_2, m_3, m_4) Using the Adoption Rate Model 99

5.4 Impressions for Small and Large Population Network 100

5.5 Parameter Estimation along Iterations ($\varepsilon, \alpha, \beta, \rho, \gamma, \omega_1, \omega_2, \omega_3, \omega_4$) for Small Population Network 102

5.6 Small Network Adoption Rate: Comparison between Survey Data and Model Outcome 103

5.7 Four Profiles of Advertisement Input on Small Network 104

5.8 Daily Adoption Rate of of Small Network with Different Advertisement Inputs (Baseline, Case 1, Case 2, Case 3) 106

5.9 Daily Adoption Rate of of Small Network with Different Advertisement Inputs for 180 Days (Baseline, Case 1) 108

5.10 Parameter Estimation Profiles for the Large Population Network 110

5.11 Large Network Adoption Rate ($P_r = 0.45$) 111

5.12 Network Influence Utility Weights along with Rewire Probability 113

5.13 Large Network Threshold with Respect to Rewire Probability 114

5.14 Parameter Sensitivity Analysis of Adoption Utility Model (Small Network) 115

5.15 Parameter Sensitivity Analysis of Adoption Utility Model (Large Network) 116

5.16 Difference of Adoption Rates in Large Population Network Achieved by Nominal Value and Monte-Carlo Calculation under Parameter Variation (Parameter ε, Standard Deviations: 10%, 20%, 50% and 100% of Mean Value) 117

6.1 Parameter Estimation along Iterations in Advertisement Control Model (Small Network) 134

6.2 Parameter Estimation along Iterations in Advertisement Control Model (Large Network) 135

6.3 Sensitivity Analysis of Model Parameters under Four Standard Deviation Levels 136

6.4 Scaled Adoption Rate Difference Between the Nominal Value and Monte-Carlo Calculation Under Parameter Variation (Parameter φ, Standard Deviations: 10%, 20%, 50% and 100% of Mean Value) 137
6.5 Baseline Daily Adoption Rates for Small and Large Networks (the First 14 days Data from Survey, the Last 14 days Data from Model Calculation under Baseline Advertisement Impressions) 139
6.6 Convergence Graph of Impression Optimisation Using Advertisement Control for Small Population Network with Method 1 140
6.7 Daily Impressions under Different Strategies in Small Population Network 141
6.8 Adoption Rates under Different Programmes of Small Network (One Baseline Strategy and Four Design Methods, Target Level at 0.7) 142
6.9 Convergence Graph of Impression Optimisation Using Advertisement Control for Small Population Network with Method 2 143
6.10 Convergence Graph of Impression Optimisation Using Advertisement Control Model for Small Population Network with Method 3 (28 days) . 145
6.11 Daily Energy Savings of Small Network (Baseline Strategy and Four Design Methods) 146
6.12 Convergence Graph of Impression Optimisation Using Advertisement Control for Small Population Network with Method 4 148
6.13 Cost Savings of Small Network (Four Design Methods) 149
6.14 Daily Deficit/Surplus of Small Network (Four Design Methods) 149
6.15 Convergence Graph of Impression Optimisation Using Advertisement Control or Large Population Network with Method 1 152
6.16 Daily Impressions for Different Strategies in Large Population Network . 153
6.17 Adoption Rates of Four Designs for Large Population Network (Four Optimal Design Results and the Baseline Result) 154
6.18 Convergence Graph of Impression Optimisation Using Advertisement Control for Large Population Network with Method 2 155
6.19 Convergence Graph of Impression Optimisation Using Advertisement Control for Large Population Network with Method 3 156

6.20 Daily Energy Savings of Large Network (Baseline Strategy and Four Design Methods) 157
6.21 Convergence Graph of Impression Optimisation Using Advertisement Control for Large Population Network with Method 4 158
6.22 Cost Savings of Large Network (Four Design Methods) 159
6.23 Daily Balance of Large Network (Four Design Methods) 160

List of Tables

4.1 Comparison of LED and Incandescent Light Bulbs 65
4.2 EES for a Single Information Source in 1 Month 70
4.3 EES for a Single Information Source in 2 Months 71
4.4 EES for a Single Information Source in 3 Months 72
4.5 EES for a Single Information Source in 6 Months 73
4.6 EES for a Single Information Source in 12 Months 73
4.7 Comparison of EES between Best Node and the Average of 10 Randomly Selected Nodes 74
4.8 EES of Top and Bottom 10 Pairs of Information Sources within 1 Month 79
4.9 EES of Top and Bottom 10 Pairs of Information Sources within 3 Months 79
4.10 EES of Top and Bottom 10 Pairs of Information Sources within 6 Months 80
4.11 EES of Top and Bottom 10 Pairs of Information Sources within 12 Months 80
4.12 Comparison of EES with Best Combinations to the Average of 10 Random Combinations 81
4.13 Cost Benefit Analysis of Single and Dual Information Sources 82

5.1 Acceptance Levels of All Segments 98
5.2 Proportion of Personal Factors (%) and Average Values 98
5.3 Coefficients of Segments 99
5.4 Small Network Model Parameters Estimation Result 102
5.5 Daily Impression of Advertisement Practice on Small Network 105
5.6 Daily Adoption Rate of of Small Network with Different Advertisement Inputs (Baseline, Case 1, Case 2, Case 3) 107

5.7 Large Network Proportion . . . 109
5.8 Large Network Model Parameters at $P_r = 0.45$ rewire probability . . . 110
5.9 Large Network Coefficients . . . 112

6.1 Estimated Model Parameters for Small and Large networks . . . 135
6.2 Optimal Advertisement Investment of Different Methods in Small Population Network . . . 141
6.3 Daily Adoption Rate of Different Methods in Small Population Network 143
6.4 EES of Baseline Method for First 14 Days in Small Population Network 146
6.5 EES of Different Methods for Last 14 Days in Small Population Network (Cumulate from Day 15) . . . 147
6.6 Daily Balance of Different Methods in Small Population Network . . . 150
6.7 Original Impression Data in Large Population Network (from Survey) . 151
6.8 Optimal Advertisement Investment of Different Methods in Large Population Network . . . 152
6.9 Daily Adoption Rate of Different Methods in Large Population Network 154
6.10 Cumulated EES of Different Methods in Large Population Network . . . 157

Nomenclature

Greek Symbols

Symbol	Description	Unit
α/α_ω	lower threshold of the adoption rate	-
$\alpha_{ij}(t)$	information propagation factor from node i to node j at time t	-
β/β_ω	upper threshold of the adoption rate	-
γ/γ_ω	diminishing return parameter of adoption rate	-
δ	probability to transit from state I to state R	-
$\Delta\tau_{ij}$	communication decay from node i to node j	-
ε	adoption threshold	-
μ, ω	advertisement adoption rate and positive response rate to advertisement	-
μ_j	attenuation factor for receiving node j	-
π	electricity tariff	£/kWh
ρ/ρ_ω	decaying rate of the adoption rate	-
$\boldsymbol{\theta}$	vector of model parameters	-
δt	sampling period	-
σ	standard deviations in Monte Carlo calculation	-

$\tilde{\tau}$	price of impressions	£
φ	probability to transit from state I to state R in information diffusion	-
Ξ	set of source node	-
ψ	probability to transit from state S to state I in information diffusion	-
$\omega_1, \omega_2, \omega_3, \omega_4$	weightings in adoption utility	-

Roman Symbols

Symbol	**Description**	**Unit**
a, b, c, d	state of income level, family situation, age, and employment status	-
A_{ij}	edge state from node i to node j	-
a_{ij}	constant propagation coefficient from node i to node j	-
$ad(t)$	advertisement influence at time t	-
B	budget limit	£
$C(T)$	cumulated cost benefit of EEP from $t = 1$ to $t = T$	£
C_0	the cost of one piece of free trial product	£
C_T	total amount of saved cost from energy savings	£
C_{adv}	the cumulated cost of advertisement	£
C_{total}	the total cost for EEP promotion (free trial product, advertisement)	£
EES_T	total EES from the influence of advertisement	kWh
$\tilde{E}_i$	energy savings from adjacent nodes connected to node i for one day	kWh
$E(t)$	effect of impressions at time t	-
E_i	energy savings of node i for one day	kWh

$\tilde{F}^T$	total EES contributed from nodes with all path length cumulated from $t = 1$ to $t = T$	kWh
f_c	average of communication frequency per month	-
F^T	total EES contributed from all source nodes cumulated from $t = 1$ to $t = T$	kWh
F_i^T	EES from node i and its adjacent nodes cumulated from $t = 1$ to $t = T$	kWh
$h(t)$	network trend influence at time t	-
$\boldsymbol{i}_{source}$	vector of information source nodes to be selected in optimisation problem	-
$\widetilde{Ins}$	final installation number including nodes with all path length	-
Inf_i	measure of energy savings information in node i	-
Ins	total installation number from adjacent nodes connected to all source node	-
Ins_i	installation number from the adjacent nodes connected to node i	-
J	residual function	-
k_i	degree of node i	-
L_i	personal acceptance level of node i	-
m_1, m_2, m_3, m_4	weighting coefficients for income level, family situation, age, and employment status	-
N	network population size	-
N_g	programme target number of product adoption	-
$n_i(t)$	adoption status of node i at time t	-
$\boldsymbol{pdf}$	probability density function	-
P_r	rewire probability in small-world network generating	-
P_{ij}	probability of information propagation from node i to node j	-

$R(t)$	adoption rate of the network at time t	-
s_i	neighbouring network influence of node i	-
S, I, R	states – aware, unaware and adopted the energy efficiency product	-
S_0, I_0, R_0	initial values for S, I and R	-
$U_i(t)$	adoption utility of node i at time t	-
v	probability to transit from state S to state I	-
W_{ij}	weight of connection from node i to node j	-
$X(t)$	impressions (number of views) at time t	-
$y(t)$	installation number derived from survey data at time t	-
$y_R(t)$	adoption rate derived from survey data	-

Acronyms

BFT Bayesian forecasting theory. iii, 46, 47, 90, 92, 121, 127, 132, 144, 160

E-R Erdos-Renyi. 34, 36, 38, 41, 111, 113, 119

EEP energy efficiency products. ii, iii, 51, 52, 54, 57–59, 65, 68–70, 82, 83, 86–90, 92, 93, 96, 97, 99, 105, 109, 112, 118–121, 124, 127, 129–131, 138, 139, 144, 145, 160

EES expected energy savings. ii, iii, ix, 50, 52, 59, 60, 67, 69–74, 76, 78–84, 88, 89, 121, 122, 128, 129, 150, 161

GA genetic algorithm. 61, 66, 98, 128

GDS Green Deal Segmentation. 97

HVAC heating, ventilation, and air conditioning. 5, 22

IET Information Entropy Theory. 29, 52

LS least square. 61, 89, 95, 96, 98, 102, 109, 118, 128, 132, 134, 160

ML machine learning. 167, 168

pdf probability density function. 75, 115–117, 137

POET performance, operation, equipment, and technology. 19

SAP Standard Assessment Procedure. 4

SIR susceptible, infected, and recovered. 123, 124

Acronyms

W-S Watts-Strogatz. v, 40, 41, 108, 111, 118, 119, 122

Preface/Acknowledgements

I wish to express my heartfelt gratitude to my supervisor, Dr Hong Yue, whose unwavering support has been instrumental in bringing this work to fruition. Her rich knowledge and experience, valuable guidance and extreme patience have been my guiding lights throughout every phase of this project.

I am also deeply appreciative of my former supervisor, Dr Jiangfeng Zhang, whose assistance and direction were pivotal in setting the foundation for my PhD journey. His insights and recommendations provided vital assistance during the initial stages of this endeavor.

Lastly, I extend special recognition to my entire family for their steadfast support and profound understanding throughout the course of my research and the creation of this thesis. Their constant encouragement has been a driving force behind my achievements.

Chapter 1

Introduction

The residential sector consistently ranks as one of the primary consumers of energy, with historic data underscoring its impact on overall energy consumption (Poortinga et al., 2004; Sartori and Hestnes, 2007).This sector's consumption patterns have significant implications for global greenhouse gas emissions, as emphasized by numerous academic and industry studies (Lutzenhiser, 1993; Van Houwelingen and Van Raaij, 1989). In the last decade, the United Kingdom has observed a marked increase in its domestic energy consumption, a trend highlighted by government reports (Palmer and Cooper, 2013). Notably, data from 2020 reveals that the domestic sector represented a third of the nation's total energy consumption, accounting for 39,276 thousand tonnes of oil equivalent or approximately 456.78 TWh (Parker, 2021). One key observation from this period is the surge in domestic energy consumption, rivaling that of the transportation sector. This uptick can be largely attributed to the COVID-19 restrictions imposed during the year, marking a unique shift in consumption patterns not seen in over three decades (Parker, 2021).

Figure 1.1 elucidates the breakdown of UK's energy consumption by various sectors. It's worth noting that the UK's domestic energy usage has, for a considerable duration, contributed to over a quarter of the country's total consumption. This trend aligns closely with associated carbon dioxide emissions (Akbari et al., 2001; Palmer and Cooper, 2013). The historical perspective on energy consumption further illuminates this narrative. Many UK buildings predate the current understanding of the climate

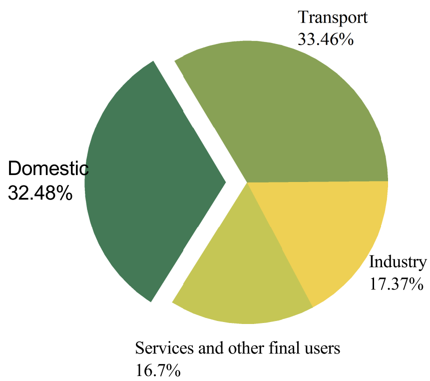


Figure 1.1: UK Energy Consumptions by Sector 2020 (Parker, 2021)

impacts linked to household energy consumption (Darby, 2006). Historically, domestic environments had lower temperatures, averaging around 12 degrees Celsius, leading inhabitants to rely on thicker clothing during winter months (Palmer and Cooper, 2013). The prevalence of central heating systems was minimal. Fast forward to today, most residences boast central heating systems, predominantly powered by gas. Additionally, the proliferation of household appliances, including refrigerators, freezers, tumble dryers, and dishwashers, further augment the domestic energy footprint. For a detailed exploration of the evolution in UK household energy consumption, readers are referred to Fig.1.2, which encapsulates data delineating these changes over time (Parker, 2021).

The accurate quantification of energy savings in residential settings necessitates a holistic consideration of multiple variables. One primary factor is weather conditions, which, coupled with the resultant changes in user behaviour, can lead to variations in energy conservation (Darby, 2006; Le Quéré et al., 2009). The report by (Palmer and Cooper, 2013) illuminates the correlation between the UK's milder winters and reduced energy requirements for heating. During the period from 1979 to 2004, heating energy consumption escalated by approximately two-fifths. This rise, however, was mitigated somewhat during phases where a concerted effort towards energy conservation was emphasized, leading to discernible reductions in energy usage (Hertin et al., 2003).

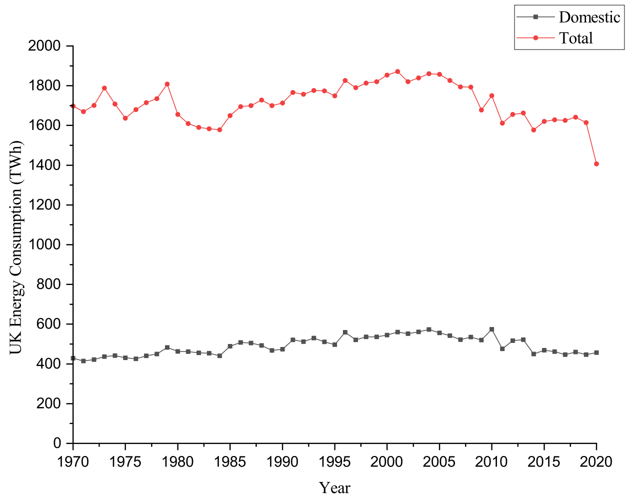


Figure 1.2: Energy Consumption of UK 2020 (Parker, 2021)

Notwithstanding these conservation efforts, studies indicate that energy saved in one area often finds its demand in other spheres (Swan and Ugursal, 2009; Xu et al., 2012). One attributable reason for this trend is the surge in population density within housing zones, which has invariably driven up the usage of basic amenities like lighting and household appliances (Wasserman, 1994).

The technological renaissance of recent decades has further compounded this demand. Innovations in electrical appliances, designed to enhance convenience and quality of life, have consequently expanded the energy footprint of most households (Palmer and Cooper, 2013). That said, advancements in housing energy efficiency, as represented by the Standard Assessment Procedure (SAP) ratings, showcase a promising trend. For context, the average SAP rating stood at 56.7 in 2012, marking a substantial improvement from the 17.6 rating observed in 1970 (Palmer and Cooper, 2013). This underscores the imperative for global energy consumption reduction, a sentiment shared not just by the residential sector but also by a diverse set of stakeholders, ranging from consumers to building energy managers (Swan and Ugursal, 2009; Wasserman, 1994). Embracing renewable energy solutions promises dual benefits: a sustainable energy model and significant reductions in carbon emissions, which, in turn, bodes well for households aiming for a reduced carbon footprint (Palmer and Cooper, 2013).

In summary, the potential energy savings achievable at the household level hold profound implications for overarching national emission goals. In this landscape, fostering proactive consumer engagement in energy conservation efforts is paramount.

1.1 Research Background

While numerous eco-projects have been initiated to foster energy-saving user behaviours (Xu et al., 2012), the early emphasis of such endeavors was predominantly on adapting to climatic conditions, inadvertently sidelining the consumers' role. Historically, the primary impetus for innovations in this sphere was climate, relegating the consumer's contribution to construction nuances and built-environment enhancement (Hertin et al., 2003). However, the tide has been turning, with consumer-centric perspectives gradually securing the spotlight in energy conservation decision-making processes. A discernible 'attitude-behaviour gap' has emerged, marked by a discrepancy between consumers' aspirational eco-friendly actions and their tangible choices.

Herein, the significance of understanding human behaviour, especially concerning energy consumption in heating, ventilation, and air conditioning (HVAC) systems, becomes paramount (Peschiera and Taylor, 2012; Xu et al., 2012). Consequently, achieving energy savings necessitates a keen exploration of consumer behaviour patterns and their alignment with sustainability principles. Today's building sector prioritizes consumer comfort, making it central to building energy management strategies. Within this context, eco-feedback systems emerge as pivotal tools, offering tangible opportunities for energy consumption reductions. These systems, when underpinned by effective energy consumption feedback mechanisms, can catalyze marked reductions in energy use. Furthermore, the role of social dynamics in energy conservation has been gaining research traction. For instance, Peschiera and Taylor (2012) employed an algorithmic methodology, rooted in empirical data from a New York City residential building, to discern energy-saving patterns. Their findings underscored the influential role of self-identified social networks in disseminating energy-saving insights.

The overarching narrative from the last decade's research is unequivocal: consumer behaviour wields considerable influence over energy savings. Simanaviciene et al. (2015)

assert the intrinsic linkage between energy savings and human behaviour, emphasizing the latter's integral role in minimizing residential energy use. Supporting this assertion, studies in (Zhang et al., 2018; Zhao et al., 2017) highlight that even in scenarios abundant with cutting-edge energy technologies, human behaviour remains the preeminent factor influencing energy savings. Delzendeh et al. (2017) further elucidate this connection, presenting correlations between energy savings and diverse behavioural factors, including income, age, and lifestyle. Yu et al. (2011) introduce the EES concept, delineating the potential energy savings achievable through the application of EEP over specific timeframes.

1.2 Problem Description

Energy consumption patterns in the residential sector are highly influenced by individual user behaviors and their social networks. While there are multiple measures for achieving energy savings, most research focuses on the technical aspects. However, quantitatively measuring and managing these influences is a less investigated area, even though it is important to have knowledge of how social behaviors, related to energy savings information, dynamically propagate. Thus, the challenge lies in the absence of practical methods and tools for quantitatively assessing residential energy savings.

1.3 Research Aims and Objectives

The aims of this project are to investigate the impacts of user behaviour on household energy savings through network systems, and design appropriate control strategies to achieve energy savings. For this purpose, the following three objectives are set.

The first objective is to develop network models that can be used to quantify the influence of user behaviour on household energy savings. Such models are required to support simulation of dynamic information diffusion, EES calculation, system analysis and optimization design.

The second objective is to examine the influence of various factors on household energy savings. This analysis will encompass individual household attributes, social network considerations, and typical scenarios related to energy conservation. Moreover, the role of advertisements as external influence on energy savings will also be scrutinized.

The third objective is to develop a controlling scheme on the most influential and manageable factor to achieve the best energy savings performance. Optimization design based on network models will be developed considering objectives under practical scenarios. By doing so, it is aspired to develop a better understanding on how to shape and and enhance energy-savings by controlling social network factor(s).

1.4 Research Contribution

The main contributions of the thesis are summarized in the following.

1. **Model development.** Several models have been developed as new contributions. Firstly, a small network model is established based on probability theory, and used to calculate EES through the information exchange within social networks in residential communities. This model is applied for the design of optimal energy usage in small-population networks. Secondly, a model is developed to integrate multiple factors that may contribute to EEP adoption decision. It is applied to simulate social networks in different scales. Thirdly, a model is established for advertisement control, in which the dynamics of the information diffusion is modelled following the mechanism of an epidemic model. Unknown model parameters are identified from statistics of survey data. The effect of advertisement on EEP adoption rate can be predicted using this network model, enabling the optimal design of advertisement investment for different EES objectives.
2. **Impacts analysis.** There are huge interests on how human behaviour and social networks may affect energy savings, but very few methods are available for quantitative analysis. Based on the model developed in this project, a systematic analysis has been conducted to examine the impacts of various factors such as individual acceptance level, influence from the connected

neighbours as well as the wide social network, and advertisement. Sensitivity analysis is applied to understand the impact of model uncertainty on the simulation results.

3. **Optimal advertisement control design.** Advertisement is found to play a crucial role influencing users' decision making on EEP adoption. Optimization design with different objectives is developed for the systems under different scenarios. The design tool can be used to help program organizers determine the optimal solutions for energy efficiency promotion among networked population.

4. **Population network campaign.** Two distinct social network campaigns are presented, enriching our understanding of network dynamics at different scales. The first offers an intimate perspective, detailing 40 individual nodes with meticulous inter-node relationships. In contrast, the second provides a holistic view of a vast network of one million users, without detailing individual connections. Together, these campaigns bridge the gap between micro-level intricacies and macro-level patterns, making a substantial contribution to the study of social networks.

1.5 Thesis Overview

The thesis work is presented in seven chapters. The links between these chapters and the flow of the work are illustrated in Fig. 1.3.

Chapter 2 – provides literature review of current work on human behaviours related to energy conservation. It assesses the extent to which the topic has been integrated into existing research and its effectiveness. The review also highlights gaps in current research, specifically the lack of a standardized, quantifiable method to study individual and social impacts to residential home energy savings.

Chapter 3 – provides a comprehensive review and preliminaries of network modeling techniques with a focus on social networks. Network models such as random networks and small-world networks are highlighted, indicating their key features and applications

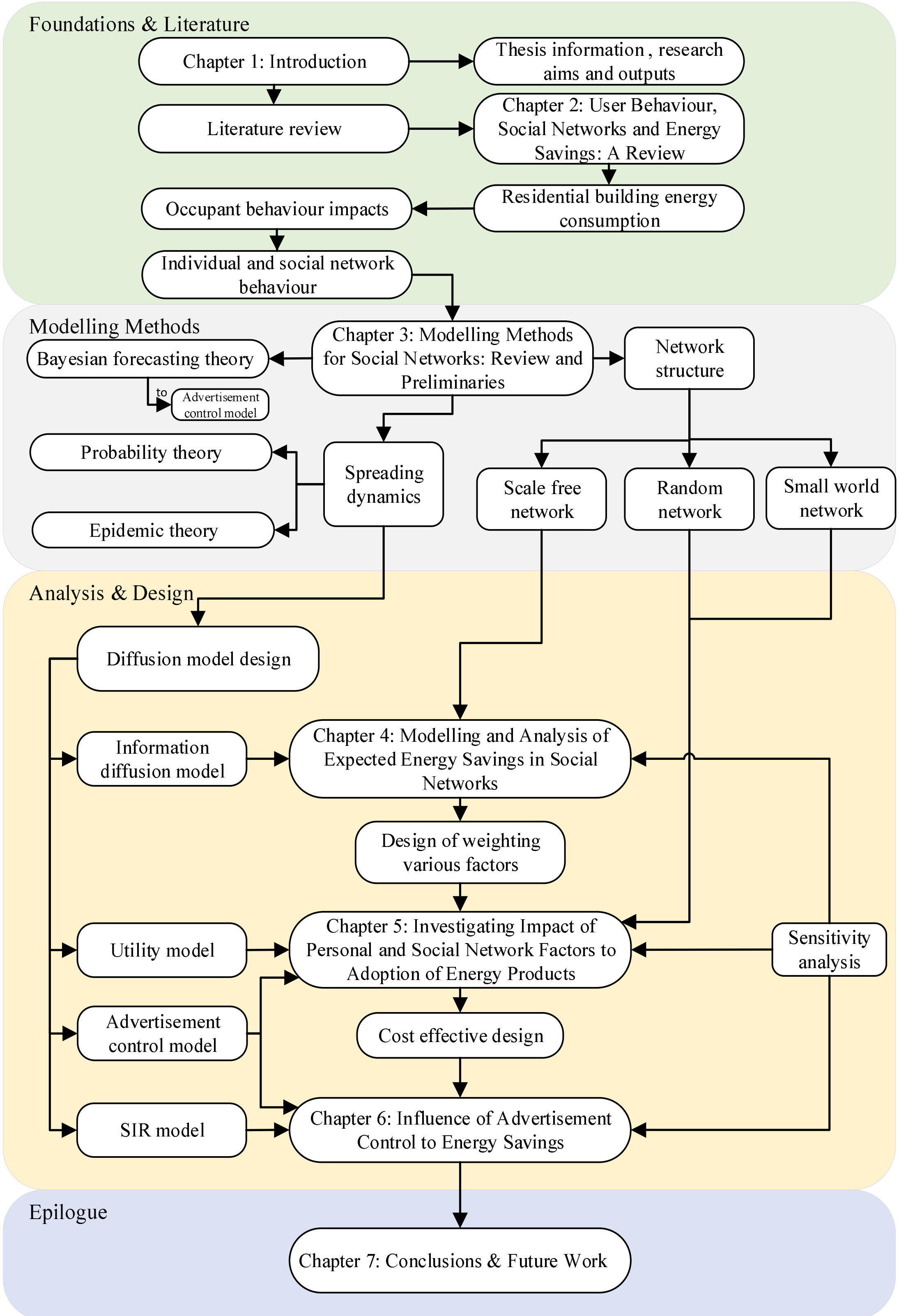


Figure 1.3: Conceptual Thesis Flow

to the latest research. An important aspect of social networks, the spreading dynamics of information within the network, is discussed. Various methods and theories applied to complex networks are analyzed, such as probability theory, epidemic theory, and BFT. These methods will be employed in Chapters 4, 5 and 6 to establish the network models required for each chapter.

Chapter 4 – presents the development of a novel approach for maximizing energy savings in small-population networks. The approach is based on probability theory in complex networks, which provides a foundation for understanding the diffusion of information through the network. Building on this foundation, a mathematical model is developed to calculate EES taking into account the diffusion of information through nodes with varying degrees of separation. This model is used to design the optimal energy usage in small-population networks. To evaluate the effectiveness of the approach, a case study is conducted using a small network of 40 participants with nodes connection coefficients. The case study provides a detailed analysis of the performance of the approach and demonstrates its potential for improving energy efficiency in small-population networks. The research result has been published in *Applied Energy* (Du et al., 2016).

Chapter 5 – presents the development of a novel approach to understand the key factors that drive the adoption of EEP among consumers. Four key factors are considered, including personal acceptance, influence from connected neighbors within the same network, overall network adoption rate, and advertisement influence. To gain a comprehensive understanding of these factors and their interplay, a network model is established, in which the four impact factors are integrated together, in a defined utility measure for EEP adoption decision making. Through the use of survey data, the relationship between adoption rate and these factors are investigated. The study indicates that advertisement plays a critical role in driving EEP adoption through a social network. The analysis results has been published at *IFAC-PapersOnline* (Du and Yue, 2023).

Chapter 6 – presents an advanced approach for maximizing EES in social networks. Using an epidemic model, the diffusion of energy-saving information within a social network is simulated. The developed model enables characterization of information diffusion, EES calculation, and optimal control of advertisement, where the change in response rate and adoption rate due to advertisement influence is analyzed using BFT. Case studies for different scenarios are conducted considering various optimization targets such as adoption rate, time cost, advertisement cost, and total energy savings subject to program budget and time constraints. The developed approach and results are published at *Applied Energy* (Du et al., 2023).

Chapter 7 – presents main conclusions of this research and offers insights into future research dimensions.

1.6 Publications

- Du, F., J. Zhang, H. Li, J. Yan, S. Galloway, and K.L. Lo, "Modelling the impact of social network on energy savings," *Applied Energy*, vol. 178, no. Supplement C, pp. 56-65, 2016. doi:10.1016/j.apenergy.2016.06.014.
  The content of the publication is included in Chapter 4.

- Du, F., H. Yue, and J. Zhang, "Influence of advertisement control to residential energy savings in large networks," *Applied Energy*, vol. 333, 120661, 2023. doi:10.1016/j.apenergy.2023.120661.
  The content of the publication is included in Chapter 5.

- Du, F. and H. Yue, "Investigating impact of individual and network factors to residential home energy savings," IFAC-PapersOnLine, vol. 56, pp. 7084-7089, 2023. doi:10.1016/j.ifacol.2023.10.573.
  The content of the publication is included in Chapter 6.

Chapter 2

User Behaviour, Social Networks and Energy Savings: A Review

2.1 Introduction

Since 1970, the rapid expansion of the UK housing market has highlighted the pressing need to examine and optimize household energy consumption (Parker, 2021). According to recent data, the UK boasts approximately 27.6 million independent housing units, commonly referred to as 'dwellings.' This number continues to grow with an annual addition of nearly 180,000 new homes (Department for Levelling Up and Communities, 2021). Paradoxically, progress in enhancing energy efficiency within these residences has been sluggish, posing a challenge to the energy sector's efforts in climate mitigation and carbon management (Fischer, 2008; Wilhite et al., 1996).

Interestingly, while the annual growth of households remains relatively stable at around 0.86%, there is a discernible trend of decreasing average household size (Fischer, 2008). This trend translates into an increased demand for individual accommodations. Regional shifts in population distribution, particularly from the North to the South West and Midlands, further influence the strategies needed for devising energy conservation methods. A modern approach to addressing this challenge revolves around the intersection of sustainable construction and social networking. These synergies serve as channels for disseminating insights on energy conservation and enhancing efficiency metrics (Brandon and Lewis, 1999; Yao and Steemers, 2005).

There is a clear push for increased transparency in household energy data. Leveraging the capabilities of sustainable architecture, stakeholders can access real-time feedback crucial for shaping energy-saving initiatives. The core objective is to promote 'smart' social buildings—structures designed to facilitate effortless sharing of energy information among residents. This concept explores the potential of social networks as essential conduits for information dissemination.

The incorporation of social networks for information dissemination is not a novel concept, but its fusion with energy conservation principles has been a recurring focus in academic investigations. This literature review aims to analyze multifaceted energy-saving strategies, centering its examination on the understanding of consumer behavior and the influential role of social networks in driving conservation efforts. The discussion encompasses the intersections of user behavior and energy consumption, modern household energy-saving frameworks, methodologies for grouping consumers into socio-groups, and the transformative possibilities of socially-driven energy-saving approaches.

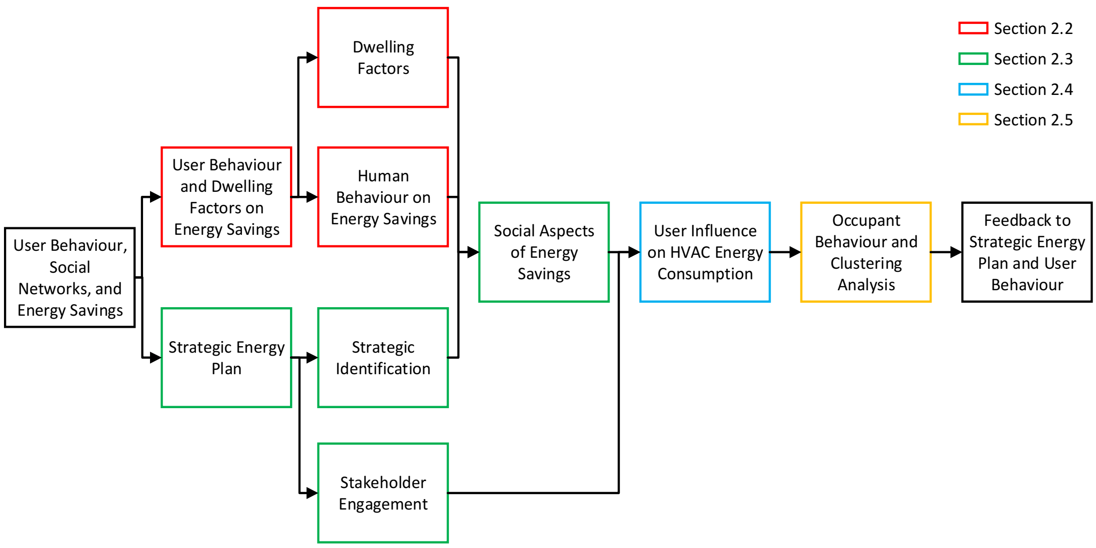


Figure 2.1: Taxonomy of User/Social Influence to Residential Home Energy

Figure 2.1 presents a network diagram corresponding to the content discussed in

Chapter 2. The section "User Behavior and Dwelling Factors on Energy Savings" emphasizes the impact of individual actions and the physical environment on energy conservation. This section is further divided into two sub-sections: "Human Behavior on Energy Savings" which directly influences the strategies outlined in the "Strategic Energy Plan," and "Dwelling Factors," highlighting the environmental and structural aspects that contribute to strategic planning for energy efficiency.

The "Strategic Energy Plan" section focuses on the significance of structured objectives and stakeholder engagement for achieving energy efficiency. It details systematic approaches to energy savings, with its subsections illustrating the planning and collaborative efforts essential for implementing energy conservation measures. Both the "User Behavior" and "Strategic Planning" aspects are influenced by and influence the "Social Aspects of Energy Savings," demonstrating the role of social dynamics and network effects in energy conservation efforts.

"User Influence on HVAC Energy Consumption" and "Occupant Behavior and Clustering Analysis" are detailed investigations into specific areas where user behavior significantly impacts energy usage. These sections highlight the practical implications of the theoretical frameworks discussed earlier and demonstrate how detailed data analysis can inform more effective energy saving strategies. There is a feedback loop from "HVAC Energy Consumption" and "Occupant Behavior Analysis" back to the "Strategic Energy Plan" and "User Behavior," illustrating the cyclical nature of analysis, implementation, and refinement in energy conservation efforts.

2.2 User Behaviour and Dwelling Factors on Energy Savings

2.2.1 Human Behaviour on Energy Savings

Human beings are active consumers of energy and thus hold a pivotal role in energy conservation. D'Oca et al. (2018) present a compelling argument: even seemingly simple actions like turning off lights or adjusting a thermostat play a crucial role in promoting energy efficiency. This viewpoint resonates throughout the academic community; Fabi et al. (2017) advocate for the intersection of human behavior and social dynamics as key areas for energy efficiency research. Interestingly, the connection between human behavior and energy conservation goes beyond individual actions. Social scientists have identified a link between behavioral science and energy management. There is a growing consensus that a collaborative approach, involving sociologists or anthropologists in the architectural process, could significantly enhance energy savings (Druckman and Jackson, 2008; Yohanis et al., 2008). These experts suggest that tenants can influence nearly 70% of a building's energy consumption patterns.

Verderber and Rubinstein (1984) substantiate this concept with empirical evidence from high-performance classrooms that demonstrate how advanced lighting systems, combined with active student engagement, lead to significant energy savings. This underscores a crucial narrative: technological innovation alone is not a panacea for energy efficiency. Human-centered designs, which integrate technological advancements with insights into human behavior, emerge as the benchmark for achieving optimal energy savings (Yan et al., 2015).

However, a gap often exists between building occupants and their architectural designers. Bridging this divide requires an intermediary role, filled by social scientists or behavioral experts who understand human needs and ensure that design interventions are consistently embraced by users (Augenbroe, 2004; Nguyen et al., 2014). With this nexus in place, informed decision-making fosters cumulative energy savings over time (Hopfe and Hensen, 2011; Williamson, 2010).

Yet, a crucial component of this energy conservation framework is training. Human behavior exhibits inherent variations, but training initiatives can harmonize these differences to optimize energy usage (Wolfe et al., 2014), especially in residential clusters such as close stock houses (Peng et al., 2012). Fischer (2008) argues that user training commitment should commence at the beginning of occupancy. Early-stage training instills energy-efficient habits in users, while subsequent training phases promote ongoing commitment and collaborative energy-saving efforts.

While modern architectural trends prioritize energy-efficient designs, one often underestimated factor is the potential of human behavior to further optimize energy efficiency. Even in buildings designed for optimal energy conservation, the way occupants interact with the infrastructure can significantly enhance or diminish energy efficiency. Kelly et al. (2013) emphasize the role of human behavior in relation to physical building elements. In some cases, building features such as insulation, heating, and cooling systems are designed with human-centric sustainability goals in mind. However, the human element can exert an influence on these variables. For instance, some individuals may use less heating during the winter, while others diligently turn off appliances when not in use. This underscores the idea that human behavior can modify the physical parameters of a building, thus impacting its energy conservation potential.

Unraveling the intricate web of human behavior is crucial for understanding its impact on energy efficiency. Factors such as income, age group, lifestyle, and other behavioral patterns are clearly linked to energy consumption (Xu et al., 2012). Furthermore, the building's environment, including factors such as location, climate, and materials, in conjunction with human behavior, shapes the potential for carbon emissions.

2.2.2 Dwelling Factors

Dwellings are distinctive and individual structures where energy efficiency is shaped by a combination of factors, including the building's environment, materials, installed equipment, and the behavior of its occupants. Kelly et al. (2013) developed a building stock model for the UK residential sector that incorporates various aspects of human behavior, such as income, age, and lifestyle. Using this model, they identified potential energy savings across different segments of residential housing and predicted how specific behaviors, like the 'rebound effect,' could influence these savings. The rebound effect refers to situations in which energy conservation achieved through efficiency measures may be offset by sudden increases in energy demand.

An important insight highlighted by Kelly et al. (2013) pertains to social behavior. Residents of energy-efficient buildings were observed to be more actively engaged, proactively utilizing the building's energy-saving features compared to those in less efficient homes. Additionally, as residents save more energy, they tend to reinvest those savings into more advanced energy conservation strategies. Notably, a study from France emphasized that environmentally conscious individuals were more inclined to undertake energy renovation measures, underscoring the influence of environmental concerns on energy-related decisions (Belaïd and Massié, 2023). Similarly, in a study conducted by Taneja and Mandys (2022), the primary drivers of electricity and gas expenses in the UK were shown to be influenced by factors such as household size, number of rooms, and urban location.

Renters exhibited greater sensitivity to energy price fluctuations compared to homeowners, highlighting the significance of dwelling tenure in shaping energy-related behaviors. Furthermore, research by İpek and İpek (2022) revealed that the determinants of residential space heating choices at the household level were heavily influenced by socio-economic and dwelling-related factors. With increasing levels of education, there was a notable shift in energy preferences from coal or wood to cleaner options like natural gas and electricity, suggesting that individuals with higher education levels tend to be more environmentally conscious.

In essence, the intricate interplay of human behaviour, the inherent attributes of buildings, and the broader housing stock variables are deeply intertwined, with each influencing the other.

2.3 Strategic Energy Plan

2.3.1 Definition and Significance

A Strategic Energy Plan is a structured approach aimed at formulating strategic objectives and implementing actions to achieve energy efficiency goals. It serves as a valuable tool for identifying and allocating resources to ensure the effective implementation of these strategic initiatives. Rather than making decisions arbitrarily, it is imperative to establish a blueprint for a comprehensive transformation of events, which should be subject to periodic review. These proactive guidelines are instrumental in executing purposeful actions that align with government directives. Comprehensive in scope, these plans encompass a wide range of community priorities and play a pivotal role in identifying high-impact opportunities for energy savings. The enduring nature of these plans, as underscored by Irawan et al. (2023), allows for flexibility and adaptation to evolving conditions over time. As demonstrated by Knuepfer et al. (2022), a well-crafted plan should remain relevant for more than a decade, emphasizing the importance of its longevity.

Given these considerations, the development of a strategic energy plan is paramount for all stakeholders. It is custom-tailored to the unique circumstances and needs of the situation. Below, we outline some fundamental protocols to be considered during the development of such plans.

To begin, it is imperative to initiate the energy conservation effort with a well-structured approach. This entails a comprehensive understanding of the repercussions and the underlying causes of energy wastage. Undertaking this endeavor is a formidable task that demands immediate attention. This meticulous comprehension was also deemed crucial in the European context, where the intricate interplay between socioeconomic challenges and energy efficiency measures was emphasized (Economidou et al., 2022). Furthermore, the political complexities associated with 'picking winners' in low-carbon research and innovation, as illustrated by the experiences of the SET-Plan, indicate that adopting more technology-neutral policies may alleviate conflicts and foster greater alignment with overarching objectives (Eikeland and Skjærseth, 2021). These initial steps lay the foundation for holistic and forward-looking strategies, as it is essential to identify current energy consumption patterns and potential energy-saving opportunities within the given context.

Secondly, it is crucial to acknowledge that not all consumers or building owners possess the suitable demographics for embracing energy-saving habits. Research conducted by Yu et al. (2011) identifies several primary factors influencing energy savings and consumption. Firstly, climate plays a pivotal role, as the environmental conditions in which a building is located significantly influence energy-saving behaviors. Factors such as ambient temperature, wind speed, and the presence of snowfall are key considerations. Secondly, construction aspects can have a profound impact on energy savings, with variables such as building design and materials affecting energy efficiency. Additionally, user characteristics, including socioeconomic backgrounds and preferences for specific service systems, exert a substantial influence on energy-saving behavior. Geographically-based indoor air quality requirements also factor into the energy-saving equation. Therefore, it is evident that social and economic factors are instrumental in shaping user attitudes towards energy conservation Peng et al. (2012); Xu et al. (2012).

2.3.2 Strategic Identification

A strategic identification of energy saving can be divided into direct savings and indirect savings (Ekpenyong et al., 2014) . Direct savings are those that are usually clearly measurable, in other words, being measurable and verifiable with proper quantitative techniques, usually expressed as a given metric involving energy saving analysis. The indirect savings on the other hand refers to mathematical calculations and expectations that are estimated from the direct savings, and which are influenced by social interactions within a framework or community following probability theory. Social interaction, as a quantifiable element, can be linked to energy efficiency and savings.

Energy efficiency and savings come under a unified classification in terms of performance, operation, equipment, and technology (POET). Ekpenyong et al. (2014) present an idea on how information transfer can help to achieve better indirect savings. Energy savings do not need to be the same across different housing sectors, optimal energy savings can be achieved by considering the maximum expected savings, the form of information sharing and the infrastructure of the residential community.

2.3.3 Stakeholder Engagement

In the development of a comprehensive plan, it is essential to incorporate stakeholder engagement and mapping. Co-creation, a collaborative approach highlighted in recent studies, has emerged as a critical element in strategic energy planning. It emphasizes the involvement of diverse actors, including the government, market entities, communities, and the third sector, throughout various phases such as initiation, design, and implementation (Sillak et al., 2021). Drawing inspiration from transitions observed in states like Illinois, the engagement of municipalities in energy planning, particularly in the context of urban planning, has gained prominence. This underscores the importance of considering local contexts and ensuring sustainable development (Rozhkov, 2023).

Among the most crucial stakeholders are local government jurisdictions, as they play a fundamental role in engaging key stakeholders with a range of interests. This underscores the significance of co-creation in ensuring efficient, effective, and socially accepted transition processes. The experiences in states like Illinois and Denmark demonstrate how local energy planning authorities, particularly in larger municipalities such as Chicago and Copenhagen, can be pivotal in designing and implementing future energy systems (Rozhkov, 2023). Collaborative participation of this nature can lead to transformative solutions in energy transitions (Sillak et al., 2021).

To ensure the successful execution of the plan, stakeholders must also mobilize resources and secure financial support for their activities. It is essential for stakeholders to comprehend the government's energy strategy, as this understanding forms a crucial component of a broader community-wide plan that includes a wide range of participants, such as environmental organizations, community activists, neighborhood representatives, and student representatives. The creation of a Stakeholder Matrix Template can further facilitate the development of solutions for various issues.

2.3.4 Social Aspects of Energy Savings

Understanding the social aspect of human behavior is undoubtedly at the core of residential energy savings. Much of the work on energy efficiency within the residential sector has traditionally been engineer-centric. Engineers tend to approach challenges from a technical standpoint, sometimes overlooking the human dimension. When interactive design is integrated into the construction of energy-efficient buildings, it becomes crucial to not only seek human input for construction feedback but also for the ongoing application of energy-saving principles (Moezzi and Janda, 2014; Owens and Driffill, 2008). A more well-rounded perspective on energy choices can only be achieved by engaging with fields such as sociology, anthropology, and other aspects related to people (Guy, 2006; Keirstead, 2006).

Certain energy-saving initiatives have been proposed by experts in the fields of behavior, energy, and climate, tailoring programs to different target audiences (Klein et al., 2012). Notably, a gap has been observed in He et al. (2023) between the intention to purchase energy-saving refrigeration appliances and the actual behavior of doing so. The decision to engage in energy-saving behaviors is not solely a matter of knowledge or intention but is also influenced by individual personalities. Specifically, individuals with emotional personalities tend to exhibit a greater inclination toward energy-saving behavior compared to those with rational personalities. Social environmental factors further amplify this inclination for individuals with emotional personalities (Ma and Liu, 2023). Additionally, the influence of online reviews related to energy efficiency on consumers' purchase decisions in the context of e-commerce platforms has been underscored, with sentiments and word-of-mouth related to energy efficiency significantly impacting sales (Ma et al., 2022).

Research by Du and Pan (2021) demonstrates that an increase in environmental awareness holds statistical significance for household energy savings, particularly among specific demographic groups such as rural households, male household heads, and young individuals. Some researchers emphasize the benefits of key energy-saving behaviors, while others advocate for a more comprehensive approach that incorporates insights from social science to comprehend diverse perspectives on energy consumption and saving.

Energy-saving programs can indeed be appealing to certain demographic groups, but to achieve significant reductions in energy consumption on a large scale, it is essential to reach a broader audience. Energy companies have developed a deep understanding of market segmentation, and user behavior, as observed within various social circles, can be considered as segments influenced by multiple factors (Klein et al., 2012). A comprehensive understanding of these nuanced segments can facilitate the promotion of better energy-saving practices. Researchers like Mazur-Stommen and Farley (2013) emphasize the need to leverage a variety of applications on specific platforms to achieve this goal.

Social integration in the context of energy savings is exemplified by apps like

Google's PowerMeter. Such apps motivate people to collaborate on energy savings. However, feedback indicates that there are distinct social segments among users. Some individuals choose to extensively use the app, while others only engage with certain features. Some even ignore the app altogether, and these divisions are often based on their level of technical proficiency.

Game design has been explored in the context of energy savings and social networks. For instance, the Tamagotchi Building Project, referenced in (Mazur-Stommen and Farley, 2013), treats individuals as 'beings' or 'pets' within their building environment, associating energy conservation with feelings of gratitude and warmth rather than solely logical or practical considerations. This project seeks to humanize buildings, promoting energy conservation as a compassionate choice.

While architectural innovations in this domain are still largely theoretical, some companies are building upon the social aspects introduced in research. For example, OPOWER is a residential energy savings project that operates on the premise that peer pressure is a powerful motivator (Allcott, 2011). OPOWER connects with utility companies, collects household energy use data, compares it with neighborhood patterns, and provides advice on energy savings, making energy-efficient choices necessary under peer pressure.

Bidgely has developed algorithms to collect energy-specific signatures from neighborhoods and offers energy-saving suggestions based on comparisons with similar areas, resulting in observed energy savings of up to 14 percent (Chakravarty and Gupta, 2013). Additionally, companies like Apple and Google have ventured into the home energy management market, with Google leading in smart-thermostat programs through its acquisition of Nest Labs.

2.4 User Influence on HVAC Energy Consumption

The HVAC (heating, ventilation, and air conditioning) system represents a significant energy consumption component in residential buildings. The interaction between HVAC systems and human behavior is intricate and multifaceted. Comprehensive multivariate analyses highlight the substantial influence of regional temperature dynamics on these behaviors. It is important to note that in subtropical climate regions, the demand for air conditioning during summer cannot be solely mitigated by modifying human behaviors.

Research indicates that more than half of total energy consumption is dedicated to maintaining indoor thermal comfort (Enteria and Mizutani, 2011; Pérez-Lombard et al., 2008). Various HVAC strategies have been evaluated for their effectiveness in reducing energy consumption, with savings ranging from 8% to 70% under different conditions (Vakiloroaya et al., 2014). These strategies have already found success in the hotel sector, leading to significant reductions in both energy usage and operational costs. A similar transition toward more efficient HVAC practices is imperative within the residential sector in the UK.

Achieving an optimal indoor environment quality necessitates an understanding of occupant behavior, including their preferences regarding heating, cooling, and how these choices impact energy consumption. Elements like HVAC operational duration, temperature set-points, and off-hour control are direct outcomes of human behavior, with research findings emphasizing their importance (Qin and Pan, 2020). These studies underscore that regardless of a building's location, occupant behaviors are pivotal in shaping energy consumption patterns. Additional factors influencing consumption include family size, nighttime ventilation practices, and the overall size of the residence.

Several studies have explored the synergy between intelligent building management systems and HVAC energy consumption, emphasizing the role of user behavior in optimizing energy conservation. For instance, Gul and Patidar (2015) propose strategies such as restricting building users' access to HVAC control systems to maximize energy savings, highlighting the occupants' significant influence on these systems' efficient operation. Therefore, it is crucial to reassess the occupant's role within the HVAC context, recognizing that an engaged and proactive approach from users can significantly impact energy savings related to HVAC systems. Zhang et al. (2017) offer valuable insights by recommending strategies such as adjusting occupants' thermal comfort range and optimizing cooling set-points to minimize energy usage, all of which directly relate to HVAC operation. They also advocate for mindful behaviors, such as avoiding extreme set-point fluctuations during the summer and optimizing the operational duration of HVAC systems. However, it is worth noting that some studies in this area lack system modeling to accurately reflect consumer behavior, an approach that can aid in the development of optimal energy conservation methods and the establishment of sustainable behavioral patterns within HVAC energy consumption.

2.5 Occupant Behaviour and Clustering Analysis

Occupant behavior plays a pivotal role in shaping a building's energy performance (Fabi et al., 2012; Guerra Santin, 2013). An effective strategy for understanding and analyzing energy-saving tendencies related to occupant behavior involves categorizing behavior into distinct clusters (Chicco, 2012; Chicco et al., 2006). This approach is particularly valuable for addressing the substantial uncertainties that can lead to significant disparities between actual energy consumption and simulation-based predictions (Hakansson et al., 2013).

Pan et al. (2017) propose a method to cluster occupant behavior based on energy usage patterns using the K-means clustering algorithm. Their data is derived from an existing case study conducted in Shanghai, where load pattern analysis was employed to identify daily, weekly, and seasonal clusters. Occupants were categorized based on their profession and economic status, specifically into three groups: white-collar workers, economically disadvantaged families, and affluent or younger families. This analysis shed light on the daily fluctuations in power consumption. Notably, the majority of households exhibited low electricity consumption. However, a distinct double peak in consumption was observed during morning and nighttime for a subset of households, primarily among the white-collar demographic. Within the economically disadvantaged cluster, this peak shift was more prominent on weekends, aligning with actual occupancy hours. One noteworthy aspect of this research is the identification of cluster patterns, providing a structured framework for categorizing occupant behavior.

Furthermore, the study by Pan et al. (2017) highlights that educational tools and awareness initiatives have a more substantial impact on occupant behavior than financial incentives, underscoring the importance of promoting understanding and knowledge to drive energy-saving practices.

While a building's energy efficiency is undeniably influenced by occupant behavior, accurately capturing and quantifying these dynamic interactions poses a significant challenge. Engineers often turn to simulations to estimate the impact of occupant behavior, especially when direct data on occupant-building interactions are limited. However, a notable limitation of many simulations, as highlighted in previous research (Guerra Santin, 2013; Pan et al., 2017), is their reliance on predefined behavioral patterns. This approach may not faithfully represent real-world scenarios, potentially introducing bias in the estimation of energy-saving potentials—especially when dealing with behaviors that do not conform to extreme cases.

To address this limitation and achieve a more accurate understanding of occupant behavior, establishing genuine relationships with building occupants can provide valuable insights into their habits. This, in turn, can refine and enhance the precision of energy simulations. In the study conducted by Stankovic et al. (2016), a novel approach was employed that combined smart meter data with qualitative data obtained from interviews. This methodology aimed to comprehensively understand the energy and time usage profiles of various domestic activities. By inferring activities such as cooking or watching TV and linking them to energy consumption patterns, this method offered a more holistic view of the energy implications of household behaviors.

Employing such an approach can significantly improve the alignment of simulated behaviors with real-world occupant patterns, ultimately enhancing the accuracy and reliability of energy performance analyses.

Further investigations underscore the importance of informing energy-saving policies and decision-making based on a nuanced understanding of residents' economic status (Andersen et al., 2009; Sun and Hong, 2017). Given the simultaneous influence of multiple factors on energy consumption, precisely isolating the individual behavioral impacts has been a challenging endeavor. To address data inconsistencies and extract meaningful insights, clustering analysis—a fundamental data mining technique—alongside min-max normalization has proven valuable. In a study conducted by Zhao et al. (2014), this approach was applied to measurements from residential buildings, demonstrating its capability to unveil energy consumption patterns associated with occupant behavior. These findings highlight the significance of behavioral modifications in curbing energy consumption, thereby aiding in the accurate modeling of occupant behavior in simulations.

A behavioral classification system, introduced by Van Raaij and Verhallen (1983*a*,*b*), categorizes occupant behavior into five distinct clusters: conservers, spenders, cool, warm, and average. Unlike economic-based categorizations, this classification is solely predicated on observable behavior and actions. The 'conservers' cluster tends to favor natural ventilation, avoiding energy-intensive solutions like HVAC systems, which represent a significant source of energy expenditure in residential settings (Gao and Malkawi, 2014; Mansur et al., 2014). The other categories, ranging from 'spenders' to 'average,' exhibit variations in temperature and ventilation preferences, with 'conservers' being the most energy-efficient group.

Additionally, Paauw et al. (2009) identified family arrangements as significant influencers of cluster behavior. Three pivotal factors that impact energy consciousness and savings emerged: family structure (single or couple), occupant actions, and convenience. The 'convenience' cluster exhibited a near-total disregard for energy conservation, while the 'energy-conscious' cluster reported a minimum of 15% savings. Similarly, Hakansson et al. (2013) identified clusters such as 'standard energy-conscious,' 'habit-driven,' and 'high-quality expectation individuals'. The latter two categories, characterized by specific habits and high expectations, often resulted in greater energy wastage. These cluster classifications provide invaluable frameworks for data integration and experimentation, contributing to an enhanced understanding of consumer behavior and its various nuances in the context of energy consumption.

2.6 Social Networks on Consumer Behaviour

Understanding consumer behavior and its clustering leads us to another crucial concept: the interplay between social networks and behavior (Boyd and Ellison, 2007; Ellison et al., 2007). Social networks have garnered significant interest in various research domains (Ellison et al., 2007; Kossinets and Watts, 2006), and in recent years, they have become a focal point for numerous academic communities (Zhao and Magoulès, 2012). In the context of energy conservation, the impact of social networks has been subject to discerning evaluation. These networks combine technological aspects with reflections of the prevailing local culture and ideology.

2.6.1 Motivational Frameworks within Social Networks

Harnessing social networks can significantly enhance consumer motivation, especially in the imperative task of reducing carbon emissions (Mankoff et al., 2007). Social networking platforms not only educate consumers but also provide them with cohesive feedback on the effectiveness of their energy-saving efforts. Various motivational frameworks can be employed to foster energy-saving-centric social networks. These networks can be dynamically engaging, facilitated through forums in housing initiatives, or even established as web portals curated by housing providers.

The tangible benefits of energy-saving clusters rooted in social networks include reductions in carbon dioxide emissions, optimized energy utilization, a shift towards environmentally-conscious lifestyles, and more (Mankoff et al., 2007). These networks invariably generate actionable insights (Centola, 2010; Kossinets and Watts, 2006; Stutzman, 2006). Methods employed by scholars focus on equipping consumers with actionable intelligence. By involving consumers in an informative social loop, they become informed about strategies to enhance energy efficiency (Ruth and Coelho, 2015). Consequently, consumers are at the center of these energy conservation efforts, as the information they share becomes crucial for devising effective energy-saving strategies.

Considering the growing transient economies and increased urban migration (Mabogunje, 1970; Rees and Wackernagel, 2008), urban energy consumption becomes a paramount concern. Therefore, energy conservation, supported by well-informed social frameworks, emerges as a pivotal imperative.

2.6.2 Utilizing Efficient Interactions in Energy Technology Adoption

Prominent studies, as highlighted in Yu et al. (2009), underscore the significance of effective interactions within societal structures for the widespread adoption of energy-efficient technologies. These mass introductions may encompass various domestic innovations, including water heaters, solar energy systems, and intelligent metering devices. Keeping housing communities regularly updated on energy efficiency innovations and conservation techniques, similar to the strategies introduced by Mohseni et al. (2021), empowers residents to make well-informed choices.

Methods proposed by Ulanowicz et al. (2009) provide approaches to quantify these efficient interactions, suggesting that both energy efficiency and conservation strategies can be improved through effective quantification. Utilizing the foundational principles of information theory, it becomes possible to establish a unified means of quantifying critical attributes, such as performance efficacy and reserve capacity. This enables the creation of a comprehensive metric for assessing the sustainability and robustness of a system (Masoso and Grobler, 2010). Communities equipped with such quantification tools are well-prepared to embrace and provide feedback on emerging energy-saving technologies.

As emphasized by (Ekpenyong et al., 2014), energy-efficient programs derive significant advantages from individual participation, which indirectly contributes to overall savings. An in-depth investigation conducted by (Swan and Ugursal, 2009) involving fifty-six households delved into the influence of social interactions on energy-saving behaviors. The primary objective was to identify households capable of effectively disseminating energy-saving information among their peers. These indirect savings materialize through the transfer and distribution of critical information within the social network.

The potential for effective quantification, drawn from population sample analyses, holds the promise of pioneering technologies that can benefit households across the UK. By sharing these energy-saving insights within social networks, a domino effect can occur, motivating a broader audience, including acquaintances and relatives of the initial information sharers, to embrace these technologies. Such shared insights may inspire recipients within this information network to further propagate the information or reconsider their perspectives on energy efficiency. This, in turn, fosters a more extensive discourse and engagement on the topic of energy efficiency.

2.6.3 Entropy-based Analysis in Energy Consumption

A recurring challenge that has been consistently identified is the need to precisely quantify the inherent uncertainties associated with data accumulation (Berta et al., 2010; Liang et al., 2002). A common presumption in variance-based research is the symmetric and equal impact of shared information across a social network. However, this assumption comes with inherent limitations, as in real-world situations, information dissemination rarely follows a uniform or symmetrical pattern.

Introduced by Shannon (1948), Information Entropy Theory (IET) provides a nuanced framework for addressing this uncertainty. IET offers a means to quantify, manipulate, and represent uncertainty levels. In the context of energy savings, entropy can be employed to assess the likelihood of achieving specific energy-saving outcomes by analyzing multiple interconnected nodes. Recent advancements have extended the application of entropy to the field of energy-efficient building retrofit strategies. For instance, the orthogonal experiment and entropy method have been utilized to objectively evaluate building retrofit schemes in rural areas, emphasizing the significance of sunspace factors in improving energy savings (Li et al., 2022).

Pioneering research has explored the interactions of social networks using IET, positioning it within the domain of information technology (Berta et al., 2010; Wellmann and Regenauer-Lieb, 2012). There is evidence, as presented in (Chen and Chen, 2015), that highlights the utility of entropy in enhancing urban energy consumption paradigms.

IET is employed in conjunction with minor variations in the energy-saving model. Those influenced by this approach such as Ekpenyong et al. (2014) are often referred to as 'free riders' (Weimann, 1994), signifying that they have not directly implemented energy efficiency programs but have instead been influenced by others, thereby benefiting from energy savings (Weinstein et al., 1989). Weimann (1994); Weinstein et al. (1989) suggest that users tend to adopt energy-saving methods that have already proven effective for their neighbors. Information entropy is employed to analyze these indirect savings.

It is worth noting that in the real world, the 'six degrees of separation' concept suggests that any two people in the world can be connected through a chain of six or fewer acquaintances (Kleinfeld, 2002; Travers and Milgram, 1977). An experiment conducted on a social network confirmed this, reporting that the average number of acquaintances separating any two people in the network is approximately 4-5 people (Backstrom et al., 2012). This implies that when a person is presented with information about the benefits of energy saving by someone who is a friend or a close acquaintance (the shortest path), they are more likely to pay attention and consider adopting energy-efficient practices (Kleinfeld, 2002). Conversely, if the information comes from someone more distant in their network, their level of interest in adopting energy-efficient practices may be lower.

Conditional probability plays a role in how information is transferred within these networks (Jain et al., 2013). Various quantitative case studies demonstrate that improvements in power-saving models can be achieved by considering the significance of households within the network. However, it's important to note that energy savings are not solely determined by the number of connections between individuals in the network. All of these factors contribute to the complexity of a social network model for energy-saving analysis. The research in (Ekpenyong et al., 2014, 2015) utilizes probability theory to develop an information diffusion model. This model is then used to quantify energy savings in the context of the IET.

2.7 Summary

The literature underscores human behaviour as central to residential energy conservation. While technological advancements aid energy savings, optimal benefits are realized when combined with behavioural changes. Dwelling types, housing variables, and human behaviour are intrinsically linked. Strategies range from direct to indirect savings with the latter emphasising the power of information diffusion through social network. Projects like OPOWER highlight the potency of peer interactions and collective energy-saving decisions. Research has extensively clustered occupant behaviours, considering work, economic, and familial factors, aiding in understanding electricity usage intricacies. The importance of quantifying behaviours is recurrent. Leveraging social networks can amplify energy-saving outcomes. Feedback systems on platforms motivate better practices, while theories like IET provide clarity on urban energy consumption nuances. The concept of 'free riders' emphasizes the communal benefits of shared energy-saving information.

2.7.1 Research Gap

Historically, homeowners prioritized the aesthetic and functional quality of their dwellings over energy efficiency measures. Those who own homes are perhaps more inclined to allocate resources to home improvements, e.g., upgrading kitchens or bathrooms, rather than investing in energy-efficient solutions like wall insulation or efficient boilers. Moreover, homeowners with financial constraints might struggle more to fund energy-efficient upgrades compared to local authorities. However, from the works examined in the literature review, it becomes clear that the mere ownership of a home, or any singular factor, shouldn't be a barrier to energy conservation. Residential energy-saving can be approached as a collective social endeavour. The literature signifies the centrality of consumer behaviour and the social exchange of energy-saving messages.

Upon evaluating the literature, distinct gaps in the subject area have become clear. Most of the studies demarcate specific research spaces – be it geographically, demographically, or otherwise – before embarking on interpretations. This specificity implies that the models developed from these studies lack general applicability. A comprehensive model, encompassing diverse residential sectors, appears to be absent. This limitation is precisely what the proposed research aims to address: develop a general model to characterize consumer behaviour intertwined with energy conservation within the framework of social networks. To develop such a model, this study will scrutinise interdisciplinary research on prevailing models of social networking and consumer energy conservation behaviours. These models will be dissected to understand their core features, and subsequently inform the creation of broad-based models. The newly proposed models can be used for system analysis and optimization design that were rarely discussed in previous work due to lack of suitable models.

Chapter 3

Modelling Methods for Social Networks: Review and Preliminaries

3.1 Introduction

Social network modelling encompasses a broad range of techniques to map and assess the flows of relationships between entities. These entities can range from individuals and groups to more complex entities such as organizations, URLs, computers, etc., as highlighted by Scott (2017). Developed following the principles of network and graph theory, social network modelling provides valuable tools for understanding and dissecting the complex architecture of social structures.

Within these networks, entities are represented as nodes, and the relationships or interactions between them are depicted as ties or edges. The power of social network modelling lies in its ability to offer both visual representations and mathematical analyses of human relationships, enabling a comprehensive understanding of relationship dynamics.

For social networks centred on energy-saving information, certain descriptors are frequently employed in research studies. Within a specific domain deemed a social network, each household is characterized as a node. Interactions amongst households are typically regarded as connections between these nodes, or edges, within the social network framework (Centola, 2010; Chen et al., 2014). For example, HVAC energy conservation can be modeled using a social network framework, where each household is considered a node, and the interactions between households are seen as connections or edges in the network. This method allows for the analysis of how energy-saving behaviors, knowledge, and practices are shared and spread among households. By understanding the structure and dynamics of these social interactions, strategies can be developed to effectively promote and enhance HVAC energy conservation within communities. This concept leverages the influence of social networks on individual and collective behavior, making it a potentially powerful tool in energy conservation efforts. In many applications, as exemplified in (Centola, 2010), edges in social network are bidirectional and bear equal weights, signifying that the influences stemming from interactions between any two nodes are symmetrical. Nevertheless, edges can carry weights and directions. As outlined in applications such as (Brandon and Lewis, 1999), the influence of interactions between two nodes might not always be balanced. The weight attributed to an edge can provide a quantitative measure of the influence exerted by one node upon another.

Figure 3.1 shows several network models used in this thesis, ordered following their randomness and heterogeneity in structures.

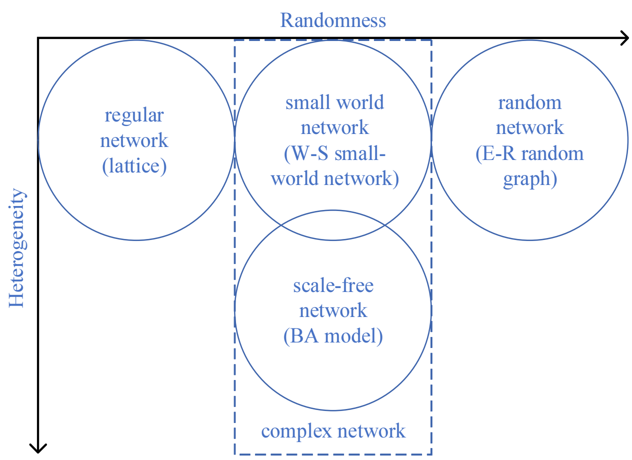


Figure 3.1: Social Network Models by Randomness and Heterogeneity

- Regular network is a network where each node has the same number of neighbors. It comprises structured clusters and is frequently used as base frameworks. Upon it, various node connection schemes can be applied to derive networks like the small-world or random networks.

- Random network is a network where each possible node pair has the same probability to be connected. It is defined by its characteristic of forming connections with equal probability, without any pattern. One thing to note is that Erdos-Renyi(E-R) random network has a fixed total number of links, thus the probability of forming connections is defined by the total number of node connections.

- Complex networks have non-trivial features that regular or random network does not have. One type of complex network is the small-world network, which has both features of clustering, like a regular network, and short path length, like a random network. In such models, the randomness of node connections lies between that of regular and random networks.

- Scale-free network is another type of complex networks characterized by the presence of a small number of highly connected nodes, often referred to as 'hubs,' while the majority of nodes have relatively few connections. In a scale-free network, the degree distribution, which represents the number of connections each node has, follows a power-law distribution.

In the following sections, these four network models will be discussed.

3.2 Regular Networks

The lattice network is a prominent example of regular networks. Clustered lattices are characterized by their tightly knit configurations, comprised of specific interaction patterns. The term 'clustered state' in this setting denotes a particular arrangement or configuration where elements (like nodes or entities) occupy a specific position or state. Often, these clusters are represented within a multi-dimensional framework known as a d-dimensional lattice (Carpineto and Romano, 1996). For $d = 2$ it can look like a typical grid with nodes connected to their nearest neighbors in horizontal and vertical directions. Figure 3.2(a) illustrates a ring lattice network with $d = 2$. Such lattices are distinct, especially given their intrinsic properties and connections that cannot be easily decoupled or separated (Meakin, 1983).

These lattices have applications in various computational and networking domains, notably in specialized computing frameworks (Stauffer, 1979). Within the broader realm of complex networks, a network where all nodes consistently exhibit the same degree of connections is classified as a regular network (Watts and Strogatz, 1998). These networks inherently display a significant degree of order and structure. A lattice network, in particular, is a deterministic regular network where each node is systematically connected to its neighboring nodes. Assuming the whole network has N nodes, the average degree of the network(average connections each node has) is k. This structure is illustrated in Fig. 3.2 (a), showcasing a regular network with ten nodes ($N = 10$) and a degree of four ($k = 4$). With $k = 4$, it can be seen that each node is connected to its four nearest neighbours (including connections on the ring). Here, P_r under the graphs represents the probability of edge re-connection, also called rewire probability.

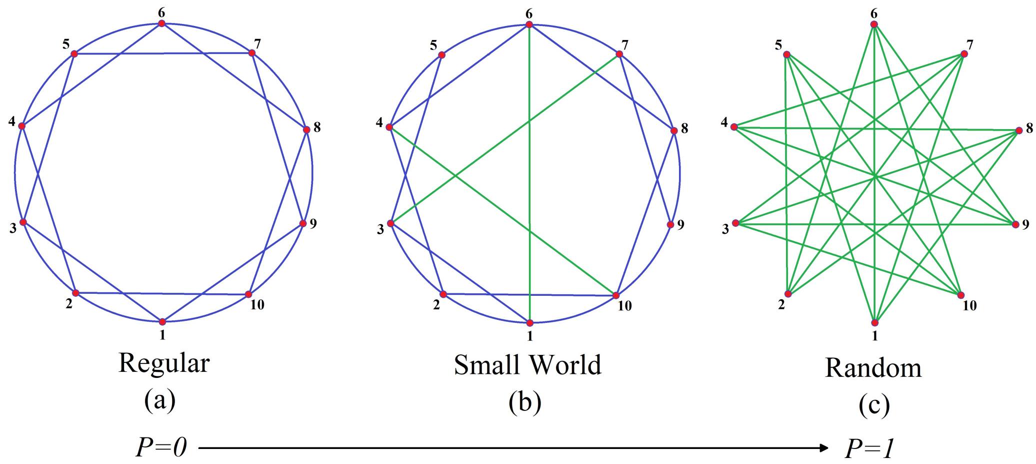


Figure 3.2: Network graphs with different randomness, from low rewire probability ($P_r = 0$) to high rewire probability ($P_r = 1$), connections before the rewiring process are in blue, after the rewired connections are in green.

3.3 Random Networks

3.3.1 Key Features and Representative Use

A random network is a theoretical construct where links are chosen randomly with equal probability. By employing a random number generator, one can establish the links from one node to another (Barabási and Albert, 1999). These random links effectively act as shortcuts to distant nodes, thereby reducing the path length. Such reduction in path length typically augments connectivity. Unlike real-world networks, random networks usually exhibit a lower degree of clustering. Moreover, they aren't suitable for representing architectural structures or the emergent patterns observed in nature. Random networks can be described either by their probability distribution or by a random process that generates them, placing the theory of random networks at the intersection of graph theory and probability theory.

Practically, random networks find their utility in scenarios where network modelling is imperative. Interestingly, a considerable proportion of these random networks mirror different types of complex networks seen in various domains (Bell and Dean, 1972). For specific type of random network, such as E-R random network, the generation of it is different from others. E-R random network includes a network with N nodes, a certain number of total edges kN are connected randomly between different two nodes, where k is the average degree of the network. Figure 3.2(c), is a showcase of E-R random network with $N = 10$ and $k = 4$.

A distinctive attribute of random graphs is percolation, indicative of a network's resilience (Callaway et al., 2000). For a random graph comprising N nodes and an average degree represented by k, removing a fraction $(1 - P_r)$ of nodes leaves a fraction of P_r nodes remained. A critical percolation threshold is determined at $P_c = 1/k$. When values are below P_c, the network undergoes fragmentation, whereas when values surpass this threshold, considerable connectivity emerges. Localised percolation involves the elimination of a node's neighbouring nodes until a fraction $(1 - P_r)$ of the total nodes are removed (Callaway et al., 2000). The random graph, characterised by a Poisson degree distribution, holds a percolation threshold of $P_c = 1/k$. Random graphs also extend their utility in probability techniques, especially when exploring the prevalence of graphs with distinct properties. The presence of a property within a random graph often indicates a broader prevalence across a majority of graphs. For a constant c, every labelled graph with n nodes and at least $cn\log(n)$ edges is Hamiltonian. Notably, certain characteristics of random graphs are contingent upon specific transformations. For instance, a transformation that equates random graph degree distributions consequently increase the clustering coefficient.

Since their inception, random networks have served as a cornerstone for groundbreaking research across diverse domains such as mathematics, engineering, social sciences, natural sciences, and medicine. Their in-depth analysis has particularly invigorated fields like discrete mathematics, probability theory, and statistical physics. One of the most intriguing phenomena associated with these networks is phase transition, which pertains to a significant change in a system's properties upon the variation of a critical parameter (Barabási et al., 2000). While phase transitions manifest across various mathematical models, they are most commonly recognized in the transitions of water: from ice to liquid and then to gas. These transitions are defined by two critical temperatures: the freezing point at zero degrees and the boiling point at one hundred degrees. At colder temperatures, atoms and molecules tend to interact closely, resulting in a structured arrangement. However, as temperatures rise, these molecular interactions wane, leading to more random configurations. A salient illustration of phase transitions in the scientific world is the theory of percolation. Rooted in physics, this theory delves into the dynamics of fluids or gases traversing porous materials. It finds relevance in diverse natural events, such as the transformation of the earth's surface due to erosion or the propagation of forest fires.

Considering a scenario where a liquid is poured atop a porous material, a prevailing question is: will the liquid be able to flow to the bottom of the porous material? This physical query is mathematically represented through bond percolation. Here, each bond, which signifies the link between two adjacent nodes in a square lattice, might be open, enabling the liquid's passage. The probability of opening a particular lattice is denoted by P, while its complement, $(1 - P)$, signifies the probability of keeping it closed, with both probabilities being mutually independent (Srinivasa and Haenggi, 2009).

3.3.2 Implications in Human Interaction and Information Propagation

Natural complexes arise from uncorrelated interactions, often forming the basis for constructed structures. Many human-engineered systems can be perceived as networks, infused with an element of randomness. Examples encompass airport and electricity networks. Random graphs are instrumental in modeling the evolution of these complex artificial structures. Intriguingly, the removal of a single node in one network can precipitate the removal of other nodes in interconnected networks. Such cascading failures can culminate in the entire network's collapse. A notable incident is the 2003 electrical blackout in Italy. The deactivation of power stations disrupted the internet network, which was integral to the power grid's management (Corsi and Sabelli, 2004).

Human movement patterns have profound implications, especially in contexts like infectious disease dissemination. While the E-R graph serves as a foundational model, it doesn't account for geographical constraints (Lusher et al., 2013). Factors like air travel introduce 'long jumps' in movement, deviating from the conventional random walks on a regular lattice. The underlying network's architecture profoundly impacts the efficacy of information or disease propagation. A central inquiry here is whether a well-connected random graph, where distant nodes are rare, is more efficient for spreading, or if a high clustering coefficient, characteristic of many real-world networks, holds greater significance. Certain human behaviors spread more rapidly and extensively within highly-clustered networks as opposed to random ones. The reinforcement from numerous neighboring nodes proves pivotal in an individual adopting and subsequently disseminating specific behaviors.

Moving to the realm of social sciences, the study of human interactions often utilizes deterministic and random graph theories. A notable illustration is the 'six degrees of separation' principle (Valente, 2010), suggesting that any two individuals are, on average, separated by a chain of six acquaintances. This bears resemblance to the E-R random graph, the human acquaintance graph exhibits a more condensed diameter and a pronounced clustering coefficient, classifying it as a small-world network which will be discussed in the following section of complex network. Thus, in such a graph, the possibility of two individuals sharing a mutual acquaintance is amplified.

3.4 Complex Networks

The advent of network science has unveiled recurrent patterns spanning diverse systems. These systems, while rooted in distinct fields like biology, economics, and physics, showcase overlapping characteristics as highlighted in (Estrada, 2011). A harmonious melding of insights from these diverse disciplines has paved the way for complex systems to be distilled and abstracted into networks or graphs. In this abstraction process, numerous details of the original system often get sidestepped, with a pronounced emphasis placed on examining the topological attributes of the complex network.

Complex networks, with their topological features, differentiate themselves from simple networks and closely mirror real-world systems (Csardi et al., 2006). The genesis of interest in these networks began with the advancement of computer research in the 20th century. A significant aspect of this interest lies in the capability to uncover and understand their complex topographical features through detailed measurements and analysis (Csardi et al., 2006). More than just abstract constructs, these networks manifest in various areas, including social and biological domains, often demonstrating distinct patterns where individual elements don't appear by mere chance.

In the vast realm of complex networks, two subcategories stand out: small-world and scale-free networks. Each has its hallmark structural attributes, with small-world networks being particularly distinct due to their characteristic high clustering while maintaining short path lengths, and scale-free network having large degree 'hub' node.

3.4.1 Small-World Networks

Small-world networks exhibit intriguing characteristics that make them stand out in the realm of complex systems. In these networks, the distance L between specific nodes tends to increase with the logarithm of the total number of nodes N. This mathematical relationship suggests that as the network grows, the average shortest path between nodes increases logarithmically (Watts and Strogatz, 1998).

Duncan Watts and Steven Strogatz, pioneering figures in network theory, discovered that small-world networks fall into a similar category as random graphs. These networks are typified not merely by their short average path length but also by an elevated clustering coefficient. Such attributes have profound implications for the dynamics and functionality of the networked systems.

When we dive into the structural details, a network or graph consists of nodes interconnected by edges. The concept of a 'path' in a network refers to a sequence of nodes and edges connecting two nodes. Interestingly, a node can feature multiple times in a specific path, giving rise to the notion of path length. In a connected graph, any node is accessible from another through a path of defined length. This accessibility brings us to the 'geodesic path,' which is the shortest path between two nodes. In the context of small-world networks, geodesic paths are notably shorter compared to other networks, lending credence to their name (Newman and Watts, 1999*a*).

Delving deeper, small-world networks frequently feature cliques or near-cliques, indicating the existence of tightly interconnected sub-networks. This characteristic contributes to the high clustering observed in small-world networks.

For a network to qualify as a small-world network, it generally possesses two key properties: a high clustering coefficient and a short average path length. The measure of a network's 'small-worldliness' can be quantified by comparing its clustering coefficient and path length to those of a comparable random network with a similar average degree (Newman and Watts, 1999*b*).

3.4.1.1 Watts-Strogatz Model: A Small-World Network Model

Duncan Watts and Steve Strogatz developed one of the most prominent small-world network models, which has deeply impacted fields such as physics, mathematics, computer science, and more. Their design, known as the Watts-Strogatz (W-S) model, elucidates the construction of a unique group of networks. These networks exhibit high clustering (transitivity) while maintaining small geodesic distances. This model generates unweighted, undirected networks devoid of self-edges or multi-edges. Such a structure bridges between two extreme configurations(regular network and random network). Here, each node connects to its nearest L(path length $\leq L$) neighbors on both sides. The resulting circular graph represents a 'large world' where it takes numerous steps to travel between any two nodes. Importantly, this configuration yields a clustering coefficient equal to $2L$, particularly when L is greater than or equal to 2 according to Watts and Strogatz (1998).

The W-S model produces networks that simultaneously exhibit pronounced clustering and the small-world phenomenon (short average path length). Delving into the mechanics of the parameter P_r, it represents the probability of an edge undergoing rewiring, where $P_r \in [0, 1]$. For $P_r = 0$, we obtain a ring graph where each node connects to its $2L$ closest neighbors, which turns out to be a regular network, shown in Fig. 3.2 (a). At $P_r = 1$, the result is an E-R random network as shown in Fig. 3.2 (c). Complex networks stay between regular networks and random networks, with $0 < P_r < 1$, as shown in Fig. 3.2 (b). This concept is visually represented in Fig. 3.2, where P_r represents the rewiring probability, connections before the rewiring process are shown in blue, while rewired connections are illustrated in green.

The W-S model, particularly when P_r is not equal to zero, is inherently probabilistic. Analyzing this model requires evaluating a collection of graphs with a constant N for varying values of P_r as pointed out in Watts and Strogatz (1998). When discussing the attributes of this model for a given P_r value, the context is usually an average over multiple graph instances. The most distinguished feature of the W-S model is the vast span of P_r values that produce small-world graphs, exhibiting pronounced clustering. This captivating characteristic has inspired a multitude of subsequent research efforts. Importantly, in the model's mechanics, new edges are introduced as shortcuts without removing any from the initial ring structure. This means that for every edge within the circular ring, there is an independent, even chance or rewire probability P_r of including a shortcut among a pair of nodes chosen uniformly or at random.

3.4.1.2 Applications of Small-World Networks

The attributes of small-world networks are not merely of academic interest; they infiltrate various real-world scenarios. Sociologically, such networks play a pivotal role in social movement groups due to their resilience to change and their efficiency in disseminating information (Newman and Watts, 1999*a*). Their structure aligns well with the idea of affinity groups: small, interdependent factions committed to a shared objective. Geophysically, the seismic networks of regions like southern California display traits of small-world networks, emphasizing their scale-invariance (Jiménez et al., 2008). In the realm of technology, these networks are instrumental in evaluating computer information usability. For instance, Freenet and Bitcoin networks showcase small-world characteristics, enhancing data retrieval efficiency for users (Newman and Watts, 1999*b*). Neurologically, the anatomical connections within the brain and the networks of cortical neurons possess small-world topologies. Such configurations, present in neuronal networks, have been linked to functions like short-term memory (Newman and Watts, 1999*a*).

3.4.2 Scale-Free Networks

A scale-free network represents a distinctive type of connected graph. The hallmark of such a network lies in its power law distribution, wherein the number of links originating from a given node is described by a power-law distribution, $P(k) \sim k^{-\gamma}$. Here $P(k)$ represents the proportion of nodes in the network with k connections to others; the exponential parameter γ typically lies between 2 and 3, although under certain situations might see it breaching these bounds (Albert and Barabási, 2002). This unique distribution suggests that some nodes have significantly more connections than others, leading to the phenomenon of hubs in the network. Constructing a scale-free network involves a dynamic and iterative process. New nodes are added to an existing network, and these nodes tend to form links with certain existing nodes. The principle guiding this linkage is the 'preferential attachment.' Essentially, a new node is more likely to form a connection with an existing node based on the number of its current connections (Albert and Barabási, 2002).

Hubs are a recurrent feature in real-world networks. These nodes, though limited in number, are highly connected. This high connectivity lends the network's degree distribution its characteristic 'long tail,' indicating the existence of nodes with degrees far exceeding the average. In the realm of scale-free networks, these hubs are not just prevalent but are also of a considerably large size, highlighting the power-law degree of distribution intrinsic to the networks.

Beyond the hubs and power-law distribution, another salient feature of scale-free networks is the degree of their nodes. Typically, nodes have a degree surpassing the average to a notable extent. The nodes with the highest degree are termed 'hubs' and their significance in the network can be vast, contingent upon the specific domain of the network. The resilience of a scale-free network, especially its robustness against failures, can be attributed to this hierarchy of hubs. Larger hubs lead the hierarchy, followed by smaller ones and subsequently by nodes of even lesser degree, as described in (Barabási et al., 2000). This structure ensures that the network remains fault-tolerant; while nodes of lower degree might be susceptible to failures, hubs, especially the major ones, are largely immune. In the rare eventuality of a hub failing, the network's intrinsic connectivity remains intact due to the presence of other hubs. However, a systematic removal of major hubs can fragment the network, a characteristic setting it apart from networks like the ER-random or small-world network.

An integral aspect of scale-free networks is the behavior of their clustering coefficient. Intriguingly, as the degree of nodes within the network increases, the clustering coefficient exhibits a decreasing trend. This kind of distribution pattern is consistent with the power rule. In practical terms, nodes with a lower degree often form part of densely interconnected sub-graphs. These sub-graphs, in turn, are often linked together through the presence of hubs.

To draw a parallel with real-world scenarios, consider social networks where individual nodes represent people and the links signify their acquaintance relationships. Within such networks, people often form tight-knit communities, resembling small clusters where every individual is acquainted with every other member. Such densely interconnected clusters can be likened to a 'complete graph.' While members of a community might have connections with individuals outside their immediate community, there are certain individuals who possess connections spanning multiple such communities; think celebrities or politicians, as discussed in Barabási (2009).

3.4.3 Comparison of Random and Complex Networks

Random graphs are typically characterized by nodes connected in a random and independent manner. Specifically, edges are incorporated between pairs of nodes based on a defined probability. On the other hand, complex networks generally originate from a set of initial nodes. As these networks evolve, new nodes tend to connect with existing ones. The probability of this connection is often proportional to the number of links that the neighboring nodes already have, as detailed in (Guimera et al., 2004). Delving deeper into their inherent structures, a complex network exhibits non-trivial topological characteristics. Such features, which set them apart from more rudimentary structures like lattices and random graphs, often align closely with graphs that model real-world systems.

Conversely, a random network, as its name suggests, is a theoretical construct where links between nodes are established randomly through probability distribution. This randomness can be achieved through the utilization of random number generators, assigning connections between nodes in a stochastic manner (Barabási and Albert, 1999).

While random networks might offer certain structural shortcuts, they generally exhibit low clustering coefficients. This makes them less suitable for real-world applications, especially in fields like computing and networking. In contrast, complex networks, which boast high clustering coefficients, are more prevalent in real-world applications due to their structural robustness and interconnectedness, as discussed in (Rubinov and Sporns, 2010).

3.5 Spreading Dynamics of Complex Networks

It's evident that epidemics, from plague, smallpox, avian flu, to SARS and the most recent COVID-19, etc., have consistently appeared in all stages of human history. While societal interconnections have bolstered public healthcare systems, reducing the threat of many epidemic diseases, this same interconnections have paradoxically amplified the potential for epidemic outbreaks due to increased human interactions.

Both information diffusion and disease spread pivot on the principle of transmission between individuals. Regardless of whether the entity in question is tangible, like a pathogen, or a piece of information, the transmission remains consistent: it typically occurs via interpersonal interaction. As this transmission proceeds, the entity permeates through a community, potentially impacting a significant portion of the population. Factors influencing the transmission's scope and speed include the entity's characteristics and the behaviors exhibited by individuals in the community.

The nature of the entity transmitted underscores a fundamental distinction between information diffusion and disease spread. While the former concerns itself with disseminating various forms of information - be it news, rumors, or advice, the latter involves the propagation of pathogens such as viruses or bacteria. These pathogens, upon infecting a host, might induce illness or other health-related repercussions.

Another salient difference lies in the transmission medium. Information predominantly disperses through communication mediums, be they verbal or written exchanges. In contrast, disease transmission often necessitates more direct methods, ranging from physical interactions to the exchange of bodily fluids.

Yet, beyond these contrasts, both processes exhibit notable similarities. The entity's dissemination from one individual to another can prompt its swift proliferation within a larger population. Furthermore, individual behaviors play crucial role in the spread, with preventive measures like protective gear against diseases or discernment against misinformation playing pivotal roles. The repercussions of both processes can be profound, with sweeping consequences for individuals and the broader society. To understand the dynamics of disease spread, researchers often employ epidemic models. These mathematical frameworks emulate the process of disease propagation within a community. They encompass various aspects, including disease transmission rates, individual susceptibility, and the efficacy of interventions like vaccinations and quarantines.

3.6 Bayesian Forecasting Theory (BFT)

The Bayesian forecasting theory (BFT) articulates that uncertainties inherent to specific situations ought to be expressed using a probability distribution. Subsequent decisions are then based on this distribution rather than solely relying on raw estimates. For instance, considering the correlation between age and the prevalence of cancer, according to the BFT, an individual's age can be harnessed to determine the probability of them having cancer. This is then compared against the probability of a person contracting cancer without factoring in their age, as described in Pole et al. (2018).

Essentially, the BFT emphasizes forecasting the probability of future events by taking into account specific influencing factors. A notable application of this theory is Bayesian inference, which presents a nuanced approach to statistical inference. When used, the probabilities associated with the BFT often contain unique nuances of interpretation. This Bayesian viewpoint guides the rational evolution of one's subjective degree of belief in response to relevant evidence. Calculating Bayesian statistics often hinges on this type of inference, as pointed out in Harrison and Stevens (1976).

The BFT is based on the use of conditional probability, defined by

$$P\left(A|B\right) = P\left(B|A\right)P\left(A\right)/P\left(B\right). \tag{3.1}$$

Here, A and B denote two events, and the condition $P(B) \neq 0$ holds as noted in Krzysztofowicz (1999). The terms in (3.1) are interpreted as follows:

- $P(A)$ and $P(B)$, the independent probabilities of observing events A and B, respectively.

- $P(A|B)$, the conditional probability, the probability of event A occurring given the occurrence of event B.

- $P(B|A)$, conditional probability, the probability of event B occurring given the occurrence of event A.

More details on BFT can be found in West and Harrison (1997).

3.7 Summary

Social networks have firmly established themselves as essential foundations in the daily functioning of our modern society. Beyond the conventional perception of social media platforms, these networks extend into crucial domains that uphold our contemporary way of life – examples include electricity grids, internet infrastructure, and communication networks.

Diving into the realm of social network modelling, it becomes apparent that the methods discussed herein hold transformative potential. These models don't just replicate real-world networks; they elucidate patterns, behaviours, and dynamics inherent within them. Through the lens of these models, nodes can be metaphorically viewed as individuals within a vast social network. Expanding on this analogy, the concept of 'node degree' can be equated to the extent of an individual's social connectivity. In this context, the adjacency matrix is no longer a mere mathematical construct; it metamorphoses into a mosaic representing social relationships and interconnections.

Exploring the realm of social network modeling reveals the transformative potential of the methods discussed here. These models do more than replicate real-world networks; they uncover patterns, behaviors, and inherent dynamics within them. Viewing nodes as individuals within a vast social network, we can expand on this analogy by equating 'node degree' with an individual's level of social connectivity. In this context, the adjacency matrix transcends being a mere mathematical construct; it transforms into a mosaic that represents social relationships and interconnections.

The essence of this review isn't just to praise the merits of network modeling but to elucidate its practical utility in capturing, comprehending, and potentially shaping social dynamics. By identifying the peculiarities of a target social network and strategically applying suitable modeling techniques, in conjunction with the insights derived from adjacency matrices, researchers can leverage this knowledge to achieve specific goals. In essence, social network modeling is not just a reflection of our interconnected world; it is a toolkit that empowers us to navigate and potentially reshape it.

This chapter reviews various graph network models in relation to their applications for understanding and influencing energy-saving behaviors through social networks. It covers regular networks, random networks, and complex networks, including small-world and scale-free networks. The chapter evaluates these models for their effectiveness in representing social interactions and information diffusion, which is crucial for promoting energy efficiency. While regular networks provide a simplistic view with evenly distributed connections, they fall short in capturing the complexity of real-world social structures. Random networks introduce randomness in connections, offering a slightly more realistic representation but still lacking in capturing the high clustering coefficient of social networks. Complex networks, especially small-world and scale-free networks, are favoured for their closer resemblance to social network structures, featuring short path lengths and a heterogeneous degree distribution, respectively. The key merit of complex network models is their ability to model clustering and influence spread effectively. By identifying the strengths and weaknesses of each network type in the context of network characterisation and effectiveness of information spread, it provides useful guide on the selection of the most appropriate modeling approaches to tackle the main problem in the thesis work.

In this research, there are two sets of campaign data with different sizes of network. The network originating from the small population campaign data closely aligns with a scale-free network, characterized by nodes with varying degrees. Notably, certain nodes demonstrate a higher number of connections to their neighbors, designating them as 'hubs'. Chapter 4 delves into an exploration of which 'hub' node or combination of nodes yields the highest EES. Chapters 5 and 6 offer a comparative analysis of the small population network against its larger counterpart.

For the large population campaign dataset, lacking detailed network connection specifics, the small-world network model becomes a fitting approach. By adjusting the re-connection probability, the network can emulate characteristics of either a regular or random network. Chapter 5 investigates the influence of diverse factors on EES within the social network. In Chapter 6, the focus shifts as data from both network sizes undergoes analysis using epidemic theory, emphasizing the propagation of EEP information from complex network connections to an overarching view of the social network.

Chapter 4

Modelling and Analysis of Expected Energy Savings in Social Networks

4.1 Introduction

In this chapter, the influence of social networks on energy saving behaviour is studied. Within a social network, individuals often share direct or indirect ties. While the influence from directly connected individuals is relatively straightforward, understanding the influence from indirect connections requires an appreciation of the small-world phenomenon. In the following sections, a model-based investigation is made to see how the interactions between individuals within social networks can potentially impact the expected energy savings (EES) of the network.

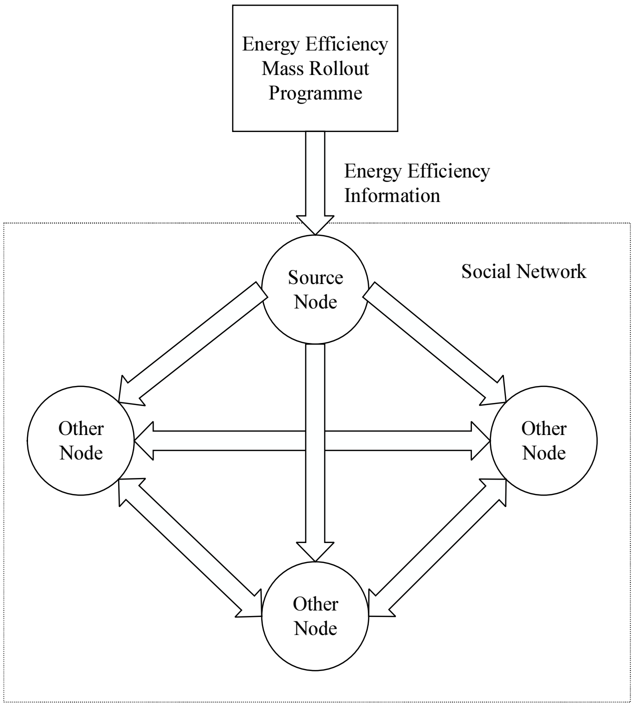


Figure 4.1: Energy Efficiency Information Spreading through Individual Interactions within a Social Network

To study information diffusion within a social network, the principles of complex network theory are employed. Within this context, an individual is represented as a node, and the interaction or connection between two individuals is taken as a node connection. Figure 4.1 illustrates the impact of interactions between individuals within a social network to the mass roll-out of energy efficiency technologies. The source node can unilaterally transmit information to nodes through direct connections; Other nodes that carry information can influence each other through their direct connection.

For instance, during the widespread implementation of EEP, such as the introduction of solar water heaters, certain nodes (individuals) within the social network adopt these efficient products ahead of others. This early adoption could be attributed to a range of reasons, such as a proactive embrace of novel technologies or the availability of free trials from vendors. Benefiting from the EEPs, these early adopters then disseminate their experiences and insights among their acquaintances, relatives, and friends. Consequently, other nodes (users) within the network become privy to this energy-saving information, potentially leading them to acquire the new product. As they experience the benefits, they too share their findings with their connections, perpetuating the spread of information. Such a ripple effect promises heightened EES, primarily driven by interpersonal interactions within the social network. However, a pertinent challenge lies in the quantitative assessment of such EES.

The intricacies of social network connections are discussed in studies like Ekpenyong et al. (2014, 2015). Here, the propagation of EES via interactions, even with minimal path length, is quantified leveraging the IET. Yet, the models in Ekpenyong et al. (2014, 2015) proceed on the assumption that each individual exerts an identical influence on all their connections, be they friends, neighbors, or family, and that this influence remains constant over time. Additionally, McCullen et al. (2013) perceive the potency of interactions between individuals as consistent and reciprocally symmetrical.

This chapter, in contrast, contemplates more practical scenarios wherein an individual's influence on their connections are different and time varying. Assisted by the framework of a weighted directed graph, a mathematical model is developed to compute the EES within a social network. Both models in (Ekpenyong et al., 2014, 2015) and this chapter utilize probability theory to quantify information diffusion and its relationship to energy savings. Instead of calculating energy savings using an entropy-based measure, the quantification of energy savings in this chapter relies on the accumulation of diffused information, calculated in the form of expected probabilities. Acknowledging the need for data for modelling, a survey is designed and implemented. This survey probes into the associations among 40 participants and gauges their reactions to EEP recommendations within the social network. Subsequent to this, the developed model undergoes validation through the data collected from the survey.

The following sections of this chapter are organized as follows. In Section 4.2, a network model is established that quantifies the EES attained via network interactions. The information source selection is presented in Section 4.3. A case study on the small network is conducted in Section 4.4, and the results are analysed. Lastly, Section 4.5 gives a summary of the chapter's findings.

4.2 Network Model Development

4.2.1 Social Networks with Connection Weightings

Within the confines of a social network, individuals are conceptualized as nodes in a complex network (Centola, 2010; Chen et al., 2014). Their mutual acquaintances or relationships, in turn, manifest as edges bridging these nodes. In this research, the strength of connections is taken into account as the influence between individuals is not assumed to be equal. The strength of connections is based on the premise that individuals who are closer to one another will have a greater impact on one another.

Consider a weighted directed network with N nodes, assume node i is connected to k_i neighbors, $i = 1, \cdots, N$. In this framework, the weight of an edge, or the connection, reflects the strength of one node's influence over the other. For the purposes of this chapter, the weight is restricted to an integer with value in the range of $[0, 5]$. A weight of zero signifies an absence of influence, indicating a lack of connection, any positive integer represents varying degrees of positive influence between nodes.

Taking into account the possibility of asymmetrical influences among individuals, the model accepts that the influence radiating from node i to node j, as measured by W_{ij}, might not reciprocate in equal measure from node j to node i, represented by W_{ji}. Consequently, the network unfolds as an unbalanced digraph, where $W_{ij} \neq W_{ji}$ may occur (Brandon and Lewis, 1999; D'Oca et al., 2018).

To quantify the influence between individuals in a social network, this research employs weights ranging from 0 to 5 for each connection. The exact value of W_{ij} depends on the frequency of communications between the two nodes (which is an assumption made by observation of participants' behaviour). This relationship is articulated as follow.

$$W_{ij} = \begin{cases} 1, & 0 \leq f_c < 1; \\ 2, & 1 \leq f_c < 2; \\ 3, & 2 \leq f_c < 4; \\ 4, & 4 \leq f_c < 10; \\ 5, & f_c \geq 10; \end{cases} \tag{4.1}$$

where f_c denotes the frequency of communications(chat between node i and j) per month. The frequency range is estimated based on individuals daily interactions. When nodes i and j are not directly connected, there will be no communications between them, thus, $W_{ij} = 0$. With these connection weights in place, we can ascertain the degree of influence between nodes.

4.2.2 Evolution of Connection Probabilities

The interactions between a pair of nodes are discussed in this section. The term $P_{ij}(t)$ refers to the probability that the information regarding an EEP is transferred from node i to its neighbour node j at time period t. This probability should encapsulate both the established connection weight and the specific details pertaining to nodes i and j, which can be expressed as

$$P_{ij}(t) = W_{ij}\alpha_{ij}(t), \tag{4.2}$$

where $\alpha_{ij}(t)$ is the propagation factor, staying between the likelihood and the weight at time t, with the constraint $0 \leq P_{ij}(t) \leq 1$. It should be noted that the formula in (4.2) is exclusive to node pairs that maintain a direct link.

Based on the epidemic theory (Chen et al., 2014; Newman, 2003), the decay of $\alpha_{ij}(t)$ can be described as a decreasing exponential function with respect to time. Hence,

$$\alpha_{ij}(t) = a_{ij}e^{-\mu_j t}, \tag{4.3}$$

where a_{ij} is the constant coefficient, which can be determined through surveys or experiments. μ_j serves as the attenuation factor for the receiving node j, given $t \geq 0$.

In this study, the effectiveness of EEP information only begins its decay after the information is received by node j from node i. Therefore, the time delay of communication must be considered in the calculation of $\alpha_{ij}(t)$.

The communication delay between two directly connected nodes, i and j, is denoted by $\Delta\tau_{ij}$. This delay is dependent on the connection strength between the pair of codes, which can be deduced through surveys or experiments tailored for a specific social network. For the small-scale social network, the communication delays are set as follow:

$$\Delta\tau_{ij} = \begin{cases} 30, \;\; W_{ij} = 1; \\ 21, \;\; W_{ij} = 2; \\ 14, \;\; W_{ij} = 3; \\ 7, \;\; W_{ij} = 4; \\ 3, \;\; W_{ij} = 5; \end{cases} \tag{4.4}$$

where the unit for $\Delta\tau_{ij}$ is measured in days in this study.

By including the decay factor with time delay, the propagation factor is revised as

$$\alpha_{ij}(t) = \begin{cases} a_{ij}e^{-\mu_j(t-\Delta\tau_{ij})}, \;\; t > \Delta\tau_{ij}; \\ 0, \;\; 0 \le t \le \Delta\tau_{ij}. \end{cases} \tag{4.5}$$

4.2.3 Information Diffusion

This subsection elucidates the comprehensive mechanism of information dissemination originating from a singular node within the network. At the outset, a node is designated as the source of information. In the initial time slot, only nodes directly linked to this primary source can receive energy saving information from it. These recipient nodes then morph into secondary source nodes, similar in essence to the originating one, and their directly linked nodes in turn become the new recipients. Figure 4.2 illustrates the information diffusion from the primary source node i within a network where the utmost path length is five. Here, nodes are classified into several groups according to their links within the network.

- i, source node;
- $q^{(1)}, q^{(2)}, q^{(3)}$, directly linked to the source node i;
- $r^{(1)}, r^{(2)}, r^{(3)}, r^{(4)}$, not directly linked to the source node i, but directly linked to nodes $q^{(1)}, q^{(2)}, q^{(3)}$;
- $s^{(1)}, s^{(2)}, s^{(3)}, s^{(4)}$, directly linked to $r^{(1)}, r^{(2)}, r^{(3)}, r^{(4)}$;
- $t^{(1)}, t^{(2)}$, directly linked to $s^{(1)}, s^{(2)}$;

- $u^{(1)}$, directly linked to $t^{(1)}$.

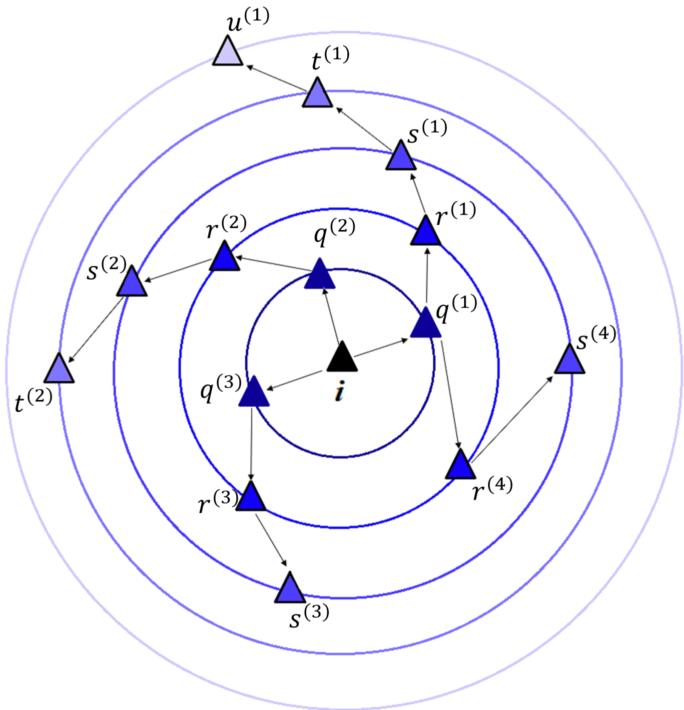


Figure 4.2: Illustration of Information Diffusion Mechanism in a Network

Starting from the source node i, it has direct links with $q^{(1)}$, $q^{(2)}$, $q^{(3)}$, facilitating a direct transfer of information from node i to each of them. Next, information are transferred from those q nodes to their connections, r nodes, and so on so forth, until the information is spread to the far most node $u^{(1)}$ from $t^{(1)}$.

This study introduces an innovative methodology to decode the dissemination of energy-saving information across social networks by juxtaposing it with the contagion of infectious diseases. In this comparative narrative, the energy-saving information is equated to a 'virus,' and individuals in the network are perceived as either 'vulnerable' or 'infected' as per epidemic theory (Chen et al., 2014). An individual can only infect others when they themselves are infected.

Different from the standard epidemic theory, in a social network, individuals may not purchase energy-saving equipment even though they are informed, they might circulate the information to their associates before they themselves adopt energy-saving products. Nonetheless, their influence wanes if they do not tangibly incorporate energy-efficient mechanisms. Consequently, influence from a peer without any tangible energy-saving setup are deemed negligible. Adhering to this assumption, it is posited that individuals can only be information propagators after they adopt the EEP. Upon adoption, a node becomes a new information source node that can further information to their connections. Hence, nodes that have a degree of separation greater than one from the nearest information source receive information indirectly, channeled through a neighbor, and only after the neighbor adopts the EEP.

Similar to the epidemiological progress, where certain individuals are immune to viruses, in a social network, there are individuals who might ignore the energy saving information. Denote these unaffected individuals as a subset G, within which nodes will not transfer information and can be ignored in the calculations of information diffusion. In a similar vein, there's another subset, denoted by K, analogous to individuals who've recovered from viral infections. In our social network context, the nodes in K cease to propagate energy-saving details after their own adoption of EEP. The sets G and K can usually be determined through dedicated surveys. The probability of information spreading from node i to node j is updated as

$$P_{ij}(t) = W_{ij}\alpha_{ij}(t), \; j \in \Pi(i), \tag{4.6}$$

where $\Pi(i)$ is the set of all the nodes that are directly connected to node i excluding those in G and K.

In theory, once nodes adopt EEP, they become new information source nodes, thereby will propagate information to their connections throughout the network. Eventually, except for those 'immune' nodes, all other nodes in the network will have adopted EEP. In practice, however, the spread of information attenuates over time and the travelling distance from the primary source i. The diminishing potency of this information, with increasing path length from the originating source, is symbolically represented by the gradually fading shades of triangles in Fig. 4.2.

Leveraging probability theory, the energy savings or the EEP installation number, through the influence from the initial source i to its adjacent nodes, can be represented in a general form as:

$$\tilde{E}_i(t) = \sum_{j \in \Pi(i)} P_{ij}(t) \cdot Inf_i. \tag{4.7}$$

The term Inf_i stands for the measure of energy savings information, in this work, it can be taken as the installation number or the energy savings contained in the initial source node i.

4.2.4 Calculation of Installation Number

For installation number calculation, define Inf_i in (4.7) as $Inf_i = 1$ for the source node i, standing for the representative installation number the source node i has, or the probability of adopting EEP is 100%. At time t, the installation number from the adjacent nodes connected to i is denoted as $Ins_i(t)$, which is

$$Ins_i(t) = \sum_{j \in \Pi(i)} P_{ij}(t) \tag{4.8}$$

Then, the installation number on the T-th day, is calculated by

$$Ins_i(T) = 1 + \sum_{t=1}^{T} Ins_i(t) = 1 + \sum_{t=1}^{T} \sum_{j \in \Pi(i)} P_{ij}(t). \tag{4.9}$$

For multiple information source nodes, the total installation number is calculated by

$$Ins(T) = \sum_{i \in \Xi} Ins_i(T) = \|\Xi\| + \sum_{i \in \Xi} \sum_{t=1}^{T} \sum_{j \in \Pi(i)} P_{ij}(t). \tag{4.10}$$

where Ξ is the set for source nodes, $\|\Xi\|$ is the number of information source nodes at the starting time point.

The above calculation of installation number is conducted for nodes adjacent to the source node i ($i \in \Xi$). Once the adjacent nodes become the new source nodes, they will influence further nodes connected to them. The information will continue to propagate until all connected nodes are covered in the network, the final installation number is denoted as $\widetilde{Ins}(T)$.

4.2.5 Calculation of Expected Energy Savings

EES can be categorized into two types: direct and indirect energy savings, depending on their origins (Ekpenyong et al., 2014). Direct savings arise from those individuals who initiate and take part in energy efficiency projects. These participants are known as information sources. On the other hand, indirect savings come from members of a social network who choose to adopt the EEP, influenced by information originally disseminated from the information sources.

In EES calculation, define Inf_i in (4.7) as E_i, standing for the energy savings of source node i for one day. At time t, the EES from the adjacent nodes to i can be calculated by

$$\tilde{E}_i(t) = \sum_{j \in \Pi(i)} P_{ij}(t) \cdot E_i. \tag{4.11}$$

The EES on the l-th day, excluding the source node contribution, is calculated by

$$\tilde{E}_i(l) = \sum_{l=1}^{l} \sum_{j \in \Pi(i)} P_{ij}(l) \cdot E_i. \tag{4.12}$$

Considering the source node contribution, the EES from the source node i and its adjacent nodes, accumulated from $t = 1$ to $t = T$ can be calculated by

$$F_i^T = \sum_{t=1}^{T} \left(E_i + \tilde{E}_i(t) \right). \tag{4.13}$$

Equation (4.13) holds true exclusively for scenarios with a single source of information within the network. For multiple resource nodes, the total EES is the aggregated contribution from all source nodes, i.e.,

$$F^T = \sum_{i \in \Xi} F_i^T. \tag{4.14}$$

It should be noted that the above calculation of EES is conducted for those nodes directly connected to node i. In a network with multi-layer connections (see Fig. 4.2), those adjacent nodes to the original source nodes will become new source nodes with their own information contents, the further EES due to the new source nodes can be calculated in a similar way, until all the connected nodes are covered in the network. The EES achieved from all indirect connections is denoted as $\tilde{F}^T$.

The pseudo-code of installation number and EES calculation is summarized in Algorithm 4.1.

Algorithm 4.1 Algorithm for Calculation of Installation Number and EES

Initialise the network connection details and state of source node, E_i, $i \in \Xi$
Update weightings of nodes W_{ij} for $i \in \Xi$, see equation (4.1)
while $0<t\leq T$ **do**
    for Each source node i **do**
        for Each receiver node j **do**
            Update propagation factor $\alpha_{ij}(t)$, see equation (4.5),
            Calculate probability $P_{ij}(t)$ see equation (4.6)
        end for
    end for
    Calculate the installation number excluding source node, $Ins_i(t)$, see equation (4.8)
    Apply the installation number calculation to all indirect nodes until the full network is covered.
    Calculate the indirect EES, $\tilde{E}_i(t)$, see equation (4.11)
    Apply the EES calculation to all indirect nodes until the full network is covered.
end while
Calculate the total installation number $Ins_i(T)$, see equation (4.9), for $i \in \Xi$.
Add the installation number from all source nodes, see equation (4.10), to get $Ins(T)$.
Apply the calculation to all the connected nodes, get $\widetilde{Ins}(T)$.
Output $\widetilde{Ins}(T)$.
Calculate the total EES F_i^T, see equation (4.13), for $i \in \Xi$.
Add the EES from all source nodes, see equation (4.14), to get F^T.
Apply the calculation to all the connected nodes, get $\tilde{F}^T$.
Output $\tilde{F}^T$.

4.2.6 Estimation of Model Parameters

Using the above model to calculate information propagation, the values of the propagation coefficient a_{ij} and the decay coefficient μ_j need to be identified. For a social network in practice, it's likely that there are unconnected nodes in the network. For modelling purpose, only those nonzero coefficients associated with connection edges need to be identified. Define the $\mathbf{a}_{ij}$ as the vector includes nonzero a_{ij}'s after vectorization of the adjacency matrix of the connections in the network. The vector of nonzero coefficients for the information propagation model is written as

$$\boldsymbol{\theta}_{\alpha} = \begin{bmatrix} \mathbf{a}_{ij}, \boldsymbol{\mu}_j \end{bmatrix}^{\top} \tag{4.15}$$

In the vector $\boldsymbol{\theta}_{\alpha} \in \mathbb{R}^n$, $\mathbf{a}_{ij}$ is the vector for nonzero a_{ij}'s, and $\boldsymbol{\mu}_j$ is the vector for nonzero μ_j's, n is the number of parameters in $\boldsymbol{\theta_{\alpha}}$.

To find out the unknown coefficients, a least square (LS) identification method is employed, which can be written as an optimization problem,

$$\boldsymbol{\theta}_{\alpha}^{*} = \arg \min_{\boldsymbol{\theta}_{\alpha} \in Z_{\boldsymbol{\theta}_{\alpha}}} J_{\boldsymbol{\theta}_{\alpha}}(\boldsymbol{\theta}_{\alpha}) \tag{4.16}$$

where $J_{\boldsymbol{\theta}_{\alpha}}$ is the residual function to be minimized. The set $Z_{\boldsymbol{\theta}_{\alpha}}$ is the searching domain for $\boldsymbol{\theta}_{\alpha}$, encompassing a set of positive real numbers. Taking the adoption data obtained from the survey over T days, the specific form of the objective function is:

$$\begin{gathered} J_{\boldsymbol{\theta}_{\alpha}} = \sum_{t=1}^{T} \left(Ins(t, \boldsymbol{\theta}_{\alpha}) - y(t)\right)^2 \\ \text{subject to } 0 < \alpha_{ij}(t) < 1 \end{gathered} \tag{4.17}$$

where $y(t)$ is the installation number of the whole network derived from survey data, which indicates the installations by time t.

The parameter identification algorithm for the information propagation model is given in Algorithm 4.2. To improve the accuracy of estimation results, the calculation proceeds in several phases. The optimization problem is solved by the genetic algorithm (GA), the results are taken as an updated initial values and refined with a gradient-based search using the Quasi-Newton method. This algorithm can be repeated for a number of times until the satisfactory results are obtained.

Algorithm 4.2 Parameter Identification Algorithm for Information Propagation Model

Initialisation
$t \leftarrow 0$
Generate population P_t with 50 feasible solutions randomly, where P_t is the population of solutions (θ_α) randomly generated (the definition of P_t applies to the whole study)
while $J_{\boldsymbol{\theta}_\alpha}(\boldsymbol{\theta}_\alpha) > 1e-6$ **do**
    Select solutions from P_t
    Recombine solutions
    Mutate solutions
    $P_{t+1} \leftarrow$ newly created solutions
    $t \leftarrow t+1$
end while
Output best solutions in P_t
Set best solution in P_t as initial value $\boldsymbol{\theta}_{\alpha 0} = [a_0, \mu_0]$, tolerance 1e-10
Interpolate the initial trial solution $\boldsymbol{\theta}_{\alpha 0}$ and calculate $J_{\boldsymbol{\theta}_\alpha}(\boldsymbol{\theta}_{\alpha 0})$
while $\frac{\partial J_{\boldsymbol{\theta}_\alpha}}{\partial \alpha} > 1e-10$ or $\frac{\partial J_{\boldsymbol{\theta}_\alpha}}{\partial \mu} > 1e-10$ **do**
    Compute the gradient of objective function $\vec{\nabla} J_{\boldsymbol{\theta}_\alpha}(\boldsymbol{\theta}_\alpha) = \left[\frac{\partial J_{\boldsymbol{\theta}_\alpha}}{\partial \alpha}, \frac{\partial J_{\boldsymbol{\theta}_\alpha}}{\partial \mu}\right]^T$. Also compute the Hessian matrix $\mathbf{H}(\boldsymbol{\theta}_\alpha) = \begin{bmatrix} \frac{\partial^2 J_{\boldsymbol{\theta}_\alpha}}{\partial \alpha^2} & \frac{\partial^2 J_{\boldsymbol{\theta}_\alpha}}{\partial \alpha \partial \mu} \\ \frac{\partial^2 J_{\boldsymbol{\theta}_\alpha}}{\partial \mu \partial \alpha} & \frac{\partial^2 J_{\boldsymbol{\theta}_\alpha}}{\partial \mu^2} \end{bmatrix}$ at point $\vec{\nabla} \boldsymbol{\theta}_{\alpha i}$
    Update the trial solution $\boldsymbol{\theta}_{\alpha(i+1)} = \boldsymbol{\theta}_{\alpha(i)} - \left[\mathbf{H}(C_{\alpha(i)})\right]^{-1} \vec{\nabla} J_{\boldsymbol{\theta}_\alpha}\left(\boldsymbol{\theta}_{\alpha(i)}\right)$
end while
Output Optimised solution

4.3 Information Sources Selection Design

In real world applications, energy saving product roll-out programme is always restricted by budget. Thus in a roll-out programme that involves free-rider (source node), the selection of the optimal source node plays a key role for promoting EEP and maximizing EES. Assume that the diffusion of energy saving information is through individual interactions only, then the proposed model can be used to simulate the whole information diffusion process, and the EES of the whole network can be calculated. In this case, optimization of the best source node for EES maximization is applicable. In this study, selection of both single and multiple information sources is designed using an optimization algorithm. The control variable is the information source number. For a design of l source nodes, denote

$$\mathbf{i_{source}} = [i_1, i_2, ..., i_l]^\top \tag{4.18}$$

where $\mathbf{i_{source}} \in \mathbb{R}^\mathbf{l}$ is the vector of l information source nodes to be selected. To achieve the maximum EES, calculated by Algorithm 4.1, the optimization problem is written as follows,

$$\mathbf{i}^*_{source} = \arg\min_{\mathbf{i} \in Z_i} -\tilde{F}^T\left(\mathbf{i_{source}}\right), \tag{4.19}$$

where Z_i is the set for all nodes in the network. The time period is the constraint of the optimization problem, as it is defined in (4.1) and (4.5), that there is a time delay in information spread between individual interactions.

4.4 Case Study of a Small Scale Social Network

4.4.1 Small Network Survey and System Setup

A survey is carried out on a group consisting of 40 users in an apartment building. Each individual is given a questionnaire that has a list of all other people's names. Participants are asked to write down the strength of relationship to other users, from zero to five, depending on the frequency of communications per month with other users.

This value of strength a participant filled in can be considered to represent the influence of other people on them. For example, if user A wrote 5 for the link strength with user B, then the weight $W_{BA} = 5$ is recorded, indicating the impact from B to A is 5.

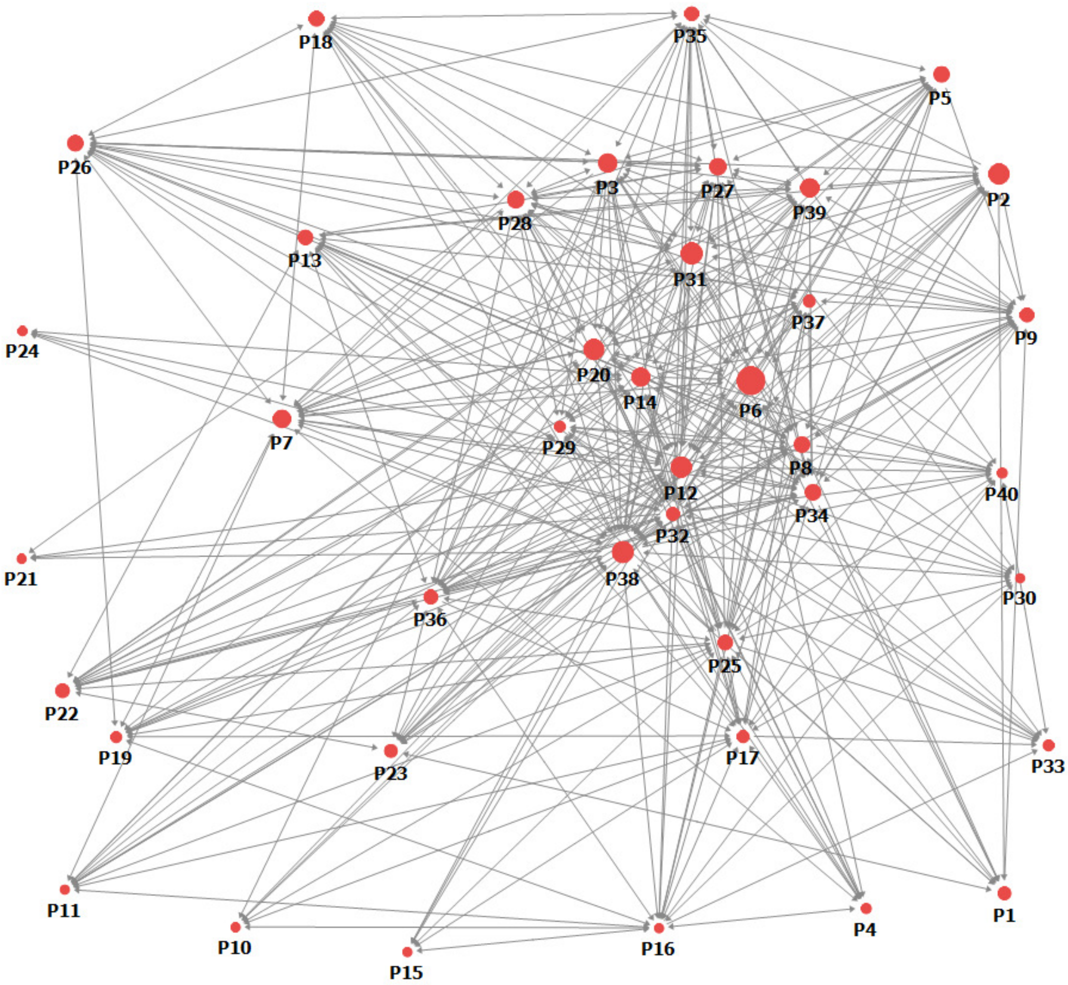


Figure 4.3: Graph Representation of the Small-Scale Social Network

In the second part of the questionnaire, questions are asked about user response to recommendations of changing incandescent light bulbs to LED light bulbs from their friends on the list. The details of the two kinds of light bulbs are shown in Table 4.1. The annual electricity bill and payback period are calculated by assuming an average of three hours' usage per day. The payback period calculated based on this table is ten months after adopting LED light bulbs. These details are listed in the questionnaire to help participants making decisions on whether to change the bulbs (electricity tariff $\pi = 0.16/kWh$).

Table 4.1: Comparison of LED and Incandescent Light Bulbs

Product Type	Price (£)	Power (W)	Lumen	Lifetime (Hours)	Annual Bill
LED	4.97	4.6	470	15,000	£ 0.81
Incandescent	1.13	40	630	2,000	£ 7.01

All the weights between nodes are obtained from the survey data, and the network connection graph is shown in Fig. 4.3. The 40 red dots with labelled numbers represent individuals in the social network. The grey lines with arrow between these dots represent directed impact between the nodes. To indicate the number of connections a node has, the sizes of the red nodes are properly scaled so that bigger size nodes have more edges connected.

Additionally, some assumptions are made for the small network survey in order to calculate EES.

- The network considers only a single piece of energy-saving information.
- Information transfer is limited to individuals with established connections.
- Every home within the building has 12 incandescent light bulbs, each used for 3 hours per day (The authors' experience is the basis for the assumed number of light bulbs and usage hours during the survey). Additionally, all these homes share the same electricity supplier.

With these assumptions, the value of E_i in (4.11) to (4.14) for this case study can be calculated according to $E_i = (p_1 - p_2) \cdot l_p \cdot 3 \cdot \Delta t$, where $l_p = 12$ indicating 12 pieces of EEP per source node, $\Delta t = 1$ indicating 1 day as time duration, $p_1 - p_2$ is the power difference between EEP and conventional product which equals to $40 - 4.6 = 35.4W$ taking data from Table 4.1. Thus, the value of E_i in this case study is $E_i = 1.2744kWh$.

4.4.2 Model Parameter Identification

With the survey data, unknown coefficients in $\boldsymbol{\theta}_\alpha$ (4.15) can be identified. For a directed network with 40 individuals, there are a total of 1560 a_{ij} and 40 μ_j to be identified. For the small network used in this case study, there are a total of 639 non-zero parameters to be identified. In the simulation implementation, two functions in Matlab, 'fmincon' and 'ga' are used. 'fmincon' employs a gradient-based search algorithm, while 'ga' utilizes the popular GA algorithm. When used individually, neither 'fmincon' nor 'ga' yields satisfactory results. Using Algorithm 4.2, the solution is first generated from 'ga' algorithm in step 1 to 11, then used as new settings of initial values for optimization using 'fmincon' algorithm from step 11 to 18. After several iterations including these two phases, the final solution yields an objective function as defined in (4.17) with a residual value of $7.2612e10^{-7}$, which is very small and acceptable considering the 639 non-zero items in the coefficient vector.

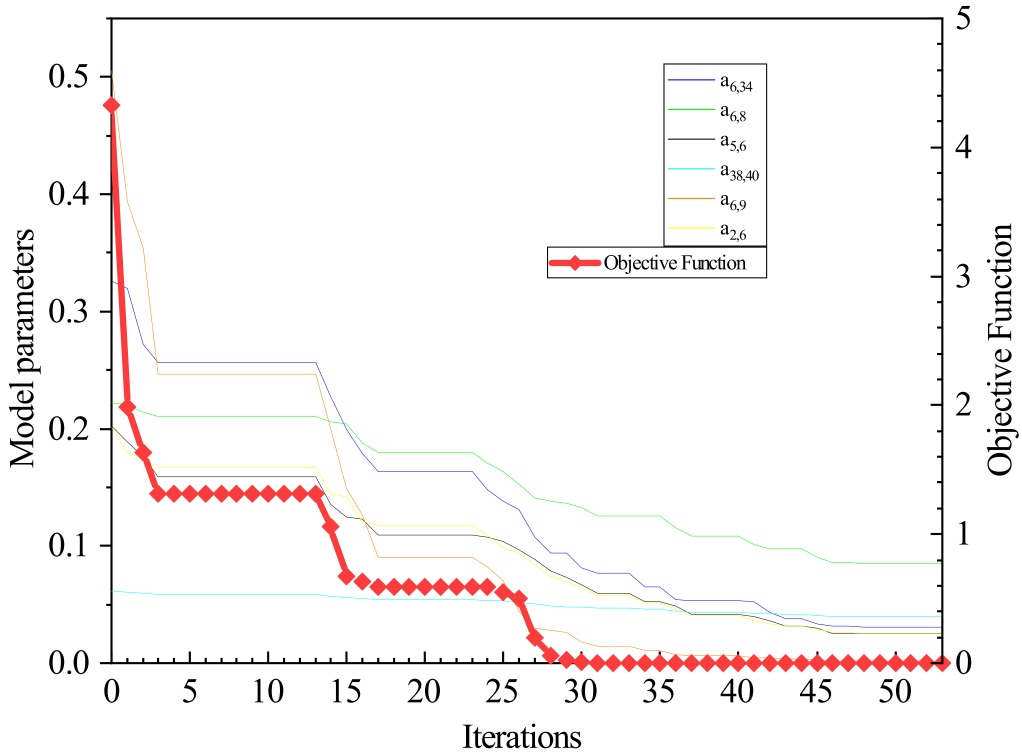


Figure 4.4: Convergence of a_{ij} Coefficients during the Estimation Process

The convergence graph of selected a_{ij}'s and μ_j's in Algorithm 4.2 are shown in Fig. 4.4 and Fig. 4.5, respectively, with corresponding objective function profiles (the red curve). Only selected parameters, according on nodes with highest degree, are shown to avoid crowded figures.

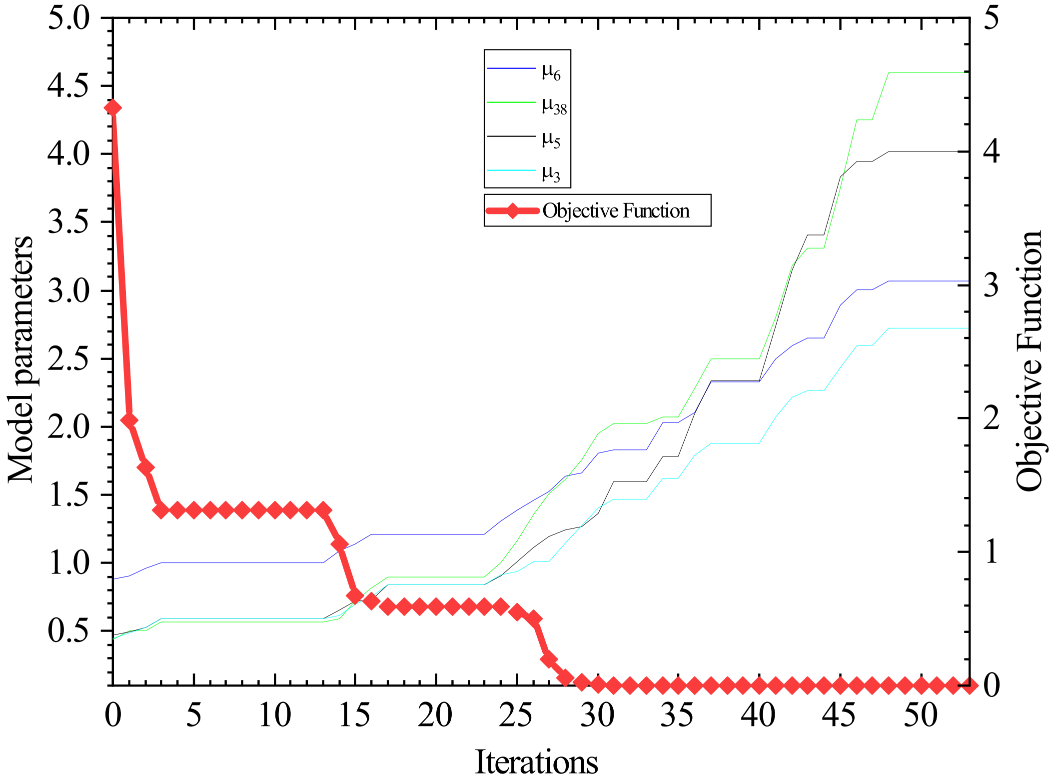


Figure 4.5: Convergence of μ_j Coefficients during the Estimation Process

4.4.3 Single Information Source

For the small social network, assume that only one person has changed all the light bulbs to LED at the beginning, that means there is a single source node. In the following discussions, two measures are used to assess energy savings performance, one is the installation number as calculated in Section 4.2.4, the other measure is the EES as modelled in Section 4.2.5.

4.4.3.1 Evolution of Installation Number

The evolution profile of installation number is an alternative measure to observe the network influence on EEP adoption. The calculation of installation number during a period of time is presented in Section 4.2.4.

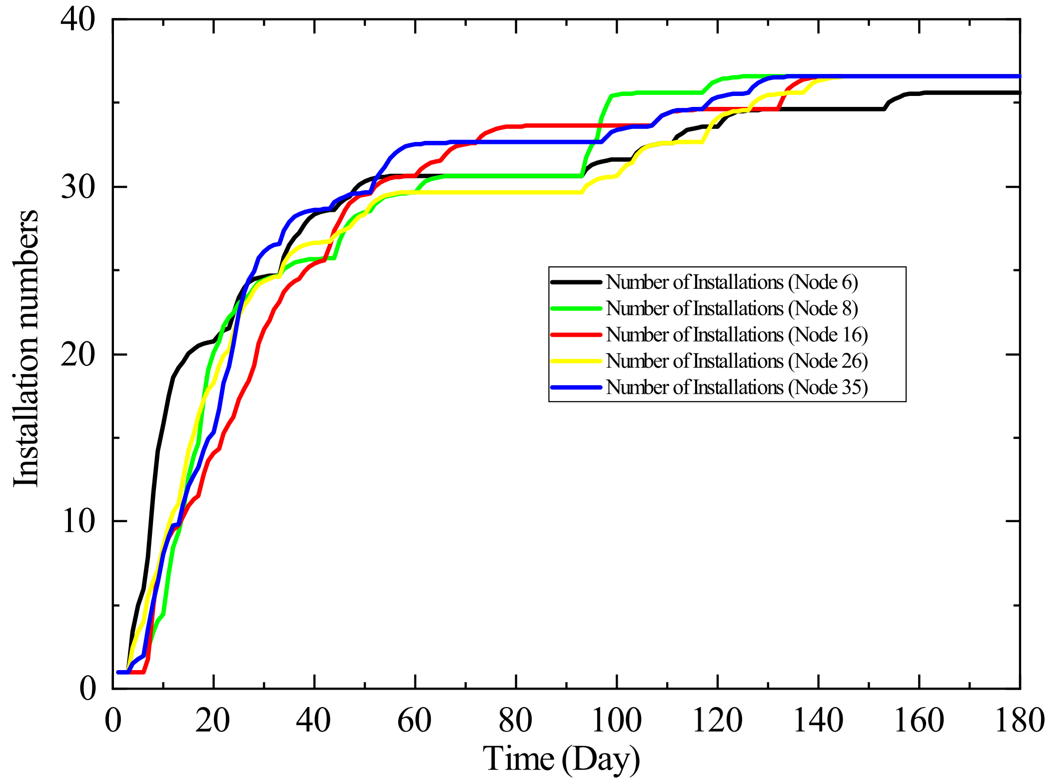


Figure 4.6: Number of Installations when Different Nodes are Selected as the Single Information Source

Figure 4.6 shows the profiles of installation number within 180 days when node 6, 8, 16, 26, 35 is taken as the source code, separately. Taking Node 6 as an example, it is observed that the installation number increases fast in the first twenty days, then slows down at the end of the first month. In the second month, the number of installations increases from 24.67 to over 30 in 20 days, and it eventually settles down at 30.65. The reason for the rapid increase of the number of installations in the initial stage is that node 6 has a large node degree ($k = 31$) and has large weighting coefficients on some of its connection edges.

The maximum value reached by node 6 in 90 days is 30.67, this value increases gradually with extended time, and eventually get saturated at 35.61 after 180 days. For all the other nodes used as single source node, it is the same that there are nearly no more increase in EEP adoption after 180 days. Node 8, 16, 26 and 35 have the highest saturated number of installations, and these values are all equal to 36.60. The other nodes are saturated at 35.61 installations. This indicates that the saturation point of a network is not highly related to the choice of the source node. For most of the source nodes within a network, they have roughly the same maximum number of installations after a sufficiently long time of period.

4.4.3.2 EES Profiles Over Different Window Lengths

Taking each user as the single source node, the calculations of EES are conducted to go through all 40 participants. The results are shown in Table 4.2 to Table 4.6, representing EES over the time duration over 1 month, 2 months, 3 months, 6 months, and 12 months, respectively. In these tables, the node degree means the number of connections a node has.

(1) Results for one month

Table 4.2: EES for a Single Information Source in 1 Month

Source Node	1	2	3	4	5	**6**	7	8
Node Degree	10	14	24	14	18	31	18	29
EES (kWh)	418.7	533.0	553.6	464.3	579.0	**644.6**	457.7	519.4
Source Node	9	10	11	12	13	14	15	16
Node degree	23	10	15	28	14	22	9	15
EES (kWh)	470.4	123.1	363.9	472.4	212.9	494.8	422.9	428.2
Source Node	17	18	19	20	21	22	23	24
Node degree	19	13	15	19	11	19	17	6
EES (kWh)	415.9	477.0	498.6	526.0	450.5	503.1	521.8	206.7
Source Node	25	26	27	28	29	30	31	32
Node degree	22	16	23	22	20	12	26	19
EES (kWh)	486.7	548.1	520.3	550.3	372.7	409.3	524.7	324.7
Source Node	33	34	35	36	37	38	39	40
Node degree	20	25	21	20	20	32	23	24
EES (kWh)	500.3	444.6	469.2	467.9	453.5	589.3	525.8	474.6

It can be seen in Table 4.2 that the largest amount of EES for the first month is 644.6 kWh when node 6 is taken as the information source. This saving is over five times larger than 123.1 kWh for which node 10 is taken as the information source. This indicates that the impact of node 6 is much larger than node 10, and in case an EEP needs to be advertised, node 6 is the most favourable target to be approached as this node will function to bring the highest possible energy savings from the whole network. Table 4.2 also shows that the adoption rate, i.e. expected savings, depends roughly but not purely on the node degree of the source node. For example, node 6 has 31 connections and node 38 has 32 connections, both are high among the 40 users and achieves higher EES when used as source node. Although node 38 has a slightly bigger node degree than node 6, its EES is lower than that of node 6. With the proposed model, the influence of a source node depends on multiple factors, the node degree is just one of them, other factors such as connection strength and communication delay also need to be considered.

Table 4.3: EES for a Single Information Source in 2 Months

Source Node	1	2	3	4	5	**6**	7	8
Node Degree	10	14	24	14	18	31	18	29
EES (kWh)	1,421.4	1,614.6	1,589.9	1,473.2	1,695.9	**1,792.3**	1,403.3	1,601.4
Source Node	9	10	11	12	13	14	15	16
Node degree	23	10	15	28	14	22	9	15
EES (kWh)	1,519.1	993.2	1,349.5	1,537.5	1,088.6	1,510.3	1,380.8	1,414.4
Source Node	17	18	19	20	21	22	23	24
Node degree	19	13	15	19	11	19	17	6
EES (kWh)	1,329.6	1,488.5	1,492.8	1,580.7	1,449.6	1,334.7	1,395.3	893.26
Source Node	25	26	27	28	29	30	31	32
Node degree	22	16	23	22	20	12	26	19
EES (kWh)	1,501.2	1,630.4	1,591.6	1,622.7	1,429.4	1,421.5	1,577.3	1,265.4
Source Node	33	34	35	36	37	38	39	40
Node degree	20	25	21	20	20	32	23	24
EES (kWh)	1,453.2	1,572.5	1,587.1	1,483.3	1,389.1	1,487.0	1,623.6	1,468.4

(2) Results for two months

Taking the time length to two months, the results of taking each user as source node are shown in Table 4.3. Again, source node 6 has the largest EES. The smallest EES occurs when taking node 24 as the source node. Node 6 achieves $1,792.3$ kWh expected savings at the end of the second month, which is much larger than its savings of 644.56 kWh at the end of the first month. This indicates that from the first month to the second month, the adoption rate within the network grows rapidly.

(3) Results for three months

Expanding the time period to three months, the results are shown in Table 4.4. It can be seen that node 6 still has the largest amount of EES, which is $2,947.7$ kWh. The incremental amount of EES for node 6 in the third month is $1,155.4$ kWh. This amount of increase is smaller than some other designated source nodes in the third month, for instance, node 5 has an increase of $1,201.6$ kWh (from $1,659.9$ kWh to $2,897.5$ kWh) in the third month. This means that the increasing rate of source node 6 is smaller than some other nodes after the second month. Due to this phenomenon, it can be seen that source node 6 has the 4th largest EES after six months as shown in Table 4.5.

Table 4.4: EES for a Single Information Source in 3 Months

Source Node	1	2	3	4	5	**6**	7	8
Node Degree	10	14	24	14	18	31	18	29
EES (kWh)	2,653.6	2,876.2	2,748.0	2,669.8	2,897.5	**2,947.7**	2,565.6	2,826.5
Source Node	9	10	11	12	13	14	15	16
Node degree	23	10	15	28	14	22	9	15
EES (kWh)	2,713.2	2,184.4	2,551.1	2,682.1	2,257.0	2,645.5	2,544.2	2,625.2
Source Node	17	18	19	20	21	22	23	24
Node degree	19	13	15	19	11	19	17	6
EES (kWh)	2,514.9	2,683.0	2,694.4	2,803.4	2,644.6	2,503.3	2,494.3	2,031.0
Source Node	25	26	27	28	29	30	31	32
Node degree	22	16	23	22	20	12	26	19
EES (kWh)	2,697.8	2,891.4	2,814.8	2,768.5	2,629.9	2,584.8	2,800.5	2,457.0
Source Node	33	34	35	36	37	38	39	40
Node degree	20	25	21	20	20	32	23	24
EES (kWh)	2,654.8	2,795.8	2,775.5	2,676.1	2,576.6	2,676.0	2,846.9	2,669.9

(4) Results for six months and 12 months

The results for 6 months and 12 months are shown in Table 4.5 and Table 4.6, respectively. For the period of 6 months, the top three nodes with large EES are 5, 35 and 8, node 6 is in the 4th place. For the period of 12 months, the top three nodes are 8, 35 and 4, node 6 drops to the 9th place as a single source. It can be seen again that the source node which has the largest amount of EES is not the node with the largest node degree. Additionally, node degrees do not seem to have any direct link with the EES. A node with higher node degree does not automatically imply it will have a higher EES. Combined factors including node degree, connection strength and communication delay will determine the influence of a source node in the network, see (4.6) for the probability of information spread between two nodes at time t.

Table 4.5: EES for a Single Information Source in 6 Months

Source Node	1	2	3	4	**5**	6	7	8
Node Degree	10	14	24	14	18	31	18	29
EES (kWh)	5,996.6	6,596.4	6,548.0	6,585.8	**6,813.5**	6,659.4	6,033.5	6,672.5
Source Node	9	10	11	12	13	14	15	16
Node degree	23	10	15	28	14	22	9	15
EES (kWh)	6,455.7	6,048.1	6,451.7	6,534.5	5,164.3	6,551.3	6,444.8	6,618.9
Source Node	17	18	19	20	21	22	23	24
Node degree	19	13	15	19	11	19	17	6
EES (kWh)	5,673.9	6,108.9	6,595.1	6,302.6	6,560.5	6,419.3	5,701.8	5,837.4
Source Node	25	26	27	28	29	30	31	32
Node degree	22	16	23	22	20	12	26	19
EES (kWh)	5,868.3	6,523.6	6,609.8	6,602.9	6,540.8	6,485.5	6,554.7	6,357.7
Source Node	33	34	35	36	37	38	39	40
Node degree	20	25	21	20	20	32	23	24
EES (kWh)	6,555.5	6,645.2	6,779.4	6,018.5	6,403.0	6,399.1	6,506.0	6,570.6

Table 4.6: EES for a Single Information Source in 12 Months

Source Node	1	2	3	4	5	6	7	**8**
Node Degree	10	14	24	14	18	31	18	29
EES (kWh)	14,203	14,694	15,092	15,461	15,229	15,310	14,696	**15,526**
Source Node	9	10	11	12	13	14	15	16
Node degree	23	10	15	28	14	22	9	15
EES (kWh)	15,071	14,876	14,898	15,151	12,063	15,416	15,123	15,443
Source Node	17	18	19	20	21	22	23	24
Node degree	19	13	15	19	11	19	17	6
EES (kWh)	12,585	14,556	15,449	15,011	14,709	15,007	14,218	9,965.7
Source Node	25	26	27	28	29	30	31	32
Node degree	22	16	23	22	20	12	26	19
EES (kWh)	14,539	14,852	15,089	15,051	14,489	15,163	15,006	14,976
Source Node	33	34	35	36	37	38	39	40
Node degree	20	25	21	20	20	32	23	24
EES (kWh)	15,415	15,155	15,466	14,272	14,948	15,268	14,994	15,056

It should be noted from Table 4.5 and Table 4.6 that the difference in the increasing rates of EES between different source nodes becomes close after the first 6 months. This is because almost all individuals in the social network, who have the probability to adopt energy saving product, have already installed the product in 6 months time.

(5) EES with best and randomly selected group of single source nodes

The source node influence to EES is further examined by comparing a best node (leading to the maximum EES) and a group of nodes selected randomly. The calculation results over a time period of one, two, three, six, and 12 months are shown in Table 4.7. In each simulation, the 10 nodes are randomly selected as the single source node, their results are averaged to give a single value to be compared with the best node. The improvement is calculated by the difference between these two over the averaged value.

In the first month, the best node achieves 39.98% more savings than the average of the 10 randomly selected nodes. This improvement figure turns to be 25.32% at the end of the second month. At the end of the 12th month, the optimal node has 5.64% more savings than the random group of nodes, and this figure turns to be smaller over time due to the gradual saturation of the EES. This result is consistent with the previous analysis in this section. It confirms the saturation pattern of EES with this small network.

Table 4.7: Comparison of EES between Best Node and the Average of 10 Randomly Selected Nodes

Months	Best node (kWh)	Ave. 10 random nodes (kWh)	Improvement
1	644.6 (node 6)	460.5	39.98%
2	1,792.3 (node 6)	1,430.2	25.32%
3	2,947.7 (node 6)	2,657.4	10.92%
6	6,813.5 (node 5)	6,398.5	6.49%
12	15,526.0 (node 8)	14,697.0	5.64%

It should be noted that the EES and the installation number are related but not the same in assessing the network influence to energy savings. In principle, a higher installation number occurred in an earlier stage of the program leads to a higher level of EES achieved over a sufficiently long period of time. If the installation number profile has a slower increase at an earlier stage, the EES achieved can be compromised even though the installation number eventually reach a high level.

4.4.4 Sensitivity Analysis with Monte Carlo Calculation

Parameter values calculated in Section 4.2.6 can be taken as nominal values for the model. The uncertainty of this model will inevitably affect the design outcome. To understand the impact from model uncertainties, sensitivity analysis and Monte Carlo calculation are performed. All parameters are assumed to follow Gaussian distribution, the mean values are the nominal values, the standard deviation is taken at four different levels, i.e., $[0.1, 0.2, 0.5, 1]$ of the mean value. The uncertainty range for each parameter is taken between zero and twice of the nominal value. In each calculation, one parameter is varied within the uncertainty range and the other parameters are kept at their nominal values.

For each parameter, nine points are uniformly sampled between the uncertainty range. Taking a_{ij} as an example, the 9 sampled points for the l_{sa}-th parameter are written as

$$a_{ij}^{l_{sa}} = [a_{ij}^{l_{sa}}(1) \cdots a_{ij}^{l_{sa}}(9)]^\top, \ \ (l_{sa} = 1, \cdots, n). \tag{4.20}$$

The values of the probability density function (pdf) corresponding to the nine sampled points are written in vector $\mathbf{pdf}_{l_{sa}}$, i.e.,

$$\mathbf{pdf}_{l_{sa}} = [\mathbf{pdf}_{l_{sa}}(1), \cdots, \mathbf{pdf}_{l_{sa}}(9)]^\top. \tag{4.21}$$

Then for the l-th sampling point in the l_{sa}-th parameter ($l = 1, \ldots, 9$), the t-th day installation number can be calculated as $Ins(t, a_{ij}^{l_{sa}}(l)) \cdot \mathbf{pdf}_{l_{sa}}(l)$.

To assess the relative impacts from each of the sampling point in a parameter, the following residual function between the calculated adoption rates and the surveyed adoption rates, weighted by pdfs of sampling points, is used as the metric for the l_{sa}-th parameter.

$$J_{l_{sa}} = \textstyle\sum_{t=1}^{T} \sum_{l=1}^{9} \Big(Ins\big(t, a_{ij}^{l_{sa}}(l)\big) - y(t) \Big)^2 \cdot \mathbf{pdf}_{l_{sa}}(l) \tag{4.22}$$

The results of $J_{l_{sa}}$ are used for sensitivity analysis.

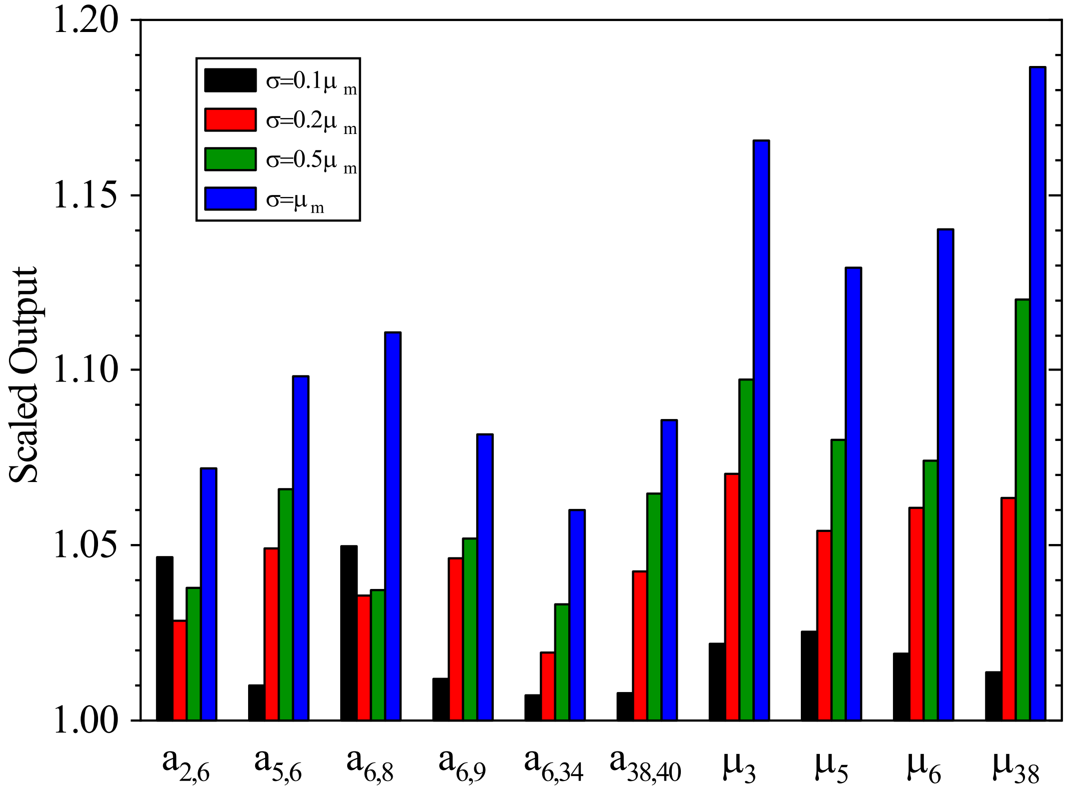


Figure 4.7: Sensitivity Analysis of EES Model Parameters under Four Standard Deviation Levels

Taking the standard deviation at four levels, i.e., 10%, 20%, 50% and 100% of the nominal value, the sensitivity analysis results for 10 selected parameters are shown in Fig. 4.7, in which the output $J_{l_{sa}}$ is normalised by the value calculated at the nominal parameters for the selected a_{ij} and μ_j.

From the sensitivity analysis, it can be observed that when the standard deviation is 10% of the nominal value, for all the 10 parameters, their uncertainty influence measure is less than 5% as calculated by (4.22). This uncertainty influence level increases when the standard deviation is increased, however, even when the standard deviation is increased to reach the nominal value, the uncertainty influence is still below 20% for all the 10 parameters. This indicates that the parametric sensitivity is low towards the model output of interest.

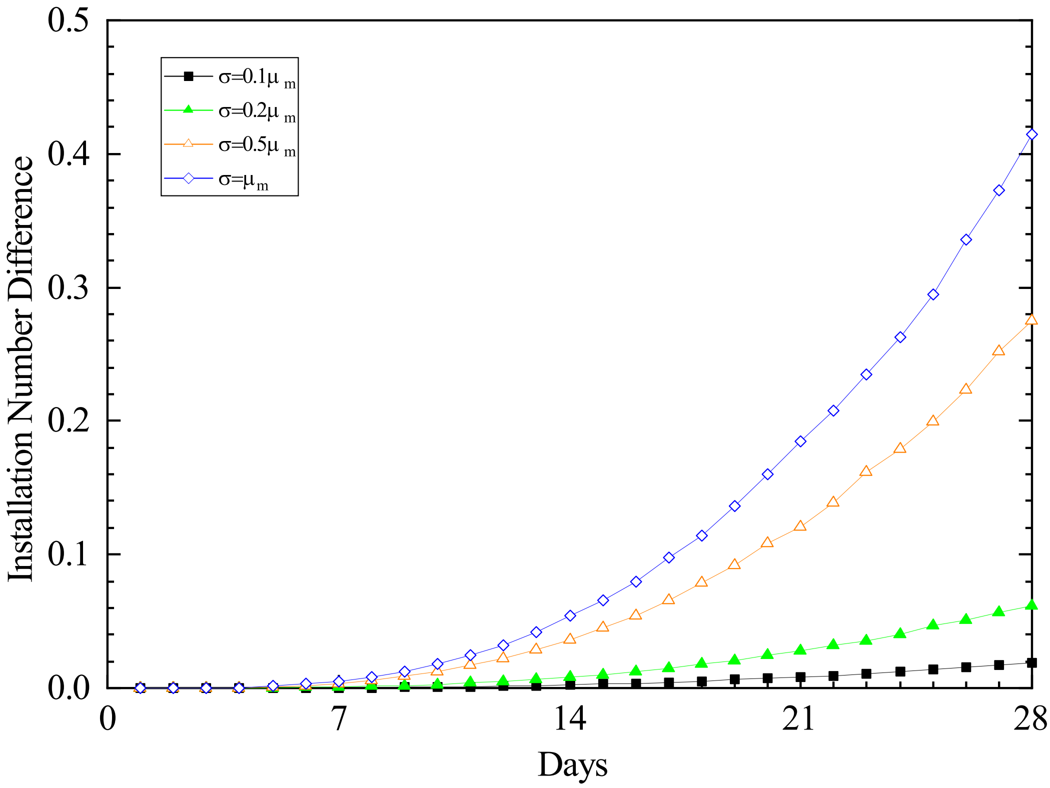


Figure 4.8: Difference of Installation Numbers Achieved by Nominal Value and Monte-Carlo Calculation under Parameter Variation(Parameter μ_{38}, Standard Deviations: 10%, 20%, 50% and 100% of Mean Value)

Among all parameters, the one that has the highest sensitivity is μ_{38}, which is taken for further analysis using the adoption rate profiles within 28 days. Again, the same sampling of nine points is applied to μ_{38}, and the calculated adoption rates are weighted by pdf values. The daily installation numbers calculated by this Monte-Carlo study are compared against the calculated installation numbers from nominal value, the difference between these two are shown in Fig. 4.8, from which it can be seen that when the parameter uncertainty is larger (larger standard deviation), the difference between the model result and the surveyed data is larger as expected. However, the difference caused by model uncertainty seems to be small, at a level of 0.5, over the 28 days. This indicates the robustness of the proposed model against parameter variations to some extent.

4.4.5 Multiple Information Sources

Now consider the case where there are two users in the network who have renovated their inefficient lighting systems into LED. In order to calculate the EES influenced by the two source nodes to the whole social network, (4.14) is applied to calculate F^T for the first-layer connections, expanded to $\tilde{F}^T$ for all the connected nodes. The values of E_i for two information sources are equal.

The EES of the whole social network is accumulated over each day during the calculation period of 1 month, 3 months, 6 months, and 12 months. In the simulation, all possible nodes pairs to be used as the information sources are calculated. The results are given in Table 4.8, Table 4.9, Table 4.10 and Table 4.11, for different time length. In these tables, the top 10 pairs (leading to large energy savings) and the bottom 10 pairs (leading to small energy savings) are listed.

(1) Results for one month and three months

In Table 4.8, the top combination is source nodes 16 and 38 which creates the maximum EES of 691.2 kWh in one month. On the other end, combination of nodes 13 and 24 produces the smallest EES in the first month with 230.3 kWh. Comparing Table 4.8 to Table 4.2, it can be seen clearly that two information sources have higher EES than a single source situation.

It is noted that although source node 6 is the best one in a single source situation, it does not appear in the best combination when two source nodes are considered. This is because that the combination of nodes 16 and 38 create more connections to the rest of the social network in comparison to the combination of node 6 with any other nodes. However, source node 6 still appears six times in the list of the 10 top combinations in Table 4.8.

Table 4.8: EES of Top and Bottom 10 Pairs of Information Sources within 1 Month

Rank	Top Pairs	EES (kWh)	Bottom Pairs	EES (kWh)
1	(16,38)	691.2	(13,24)	230.3
2	(6,17)	689.0	(10,13)	236.6
3	(5,16)	687.7	(10,24)	251.3
4	(6,34)	686.4	(10,32)	304.6
5	(6,8)	682.1	(10,11)	310.6
6	(6,16)	682.0	(1,10)	315.0
7	(5,6)	681.6	(10,30)	331.8
8	(33,38)	677.8	(10,15)	334.1
9	(6,40)	675.7	(24,32)	343.2
10	(38,40)	675.1	(1,32)	347.2

When a longer period is considered, for 3 months savings, the combination of nodes 6 and 8 is the best resulting in $3,057.1$ kWh of savings (see Table 4.9). The best combination for 1 month becomes the 7th at the end of 3 months. It is also noted that the EES difference between the top 10 combinations is very small for each time period under study.

Table 4.9: EES of Top and Bottom 10 Pairs of Information Sources within 3 Months

Rank	Top Pairs	EES (kWh)	Bottom Pairs	EES (kWh)
1	(6,8)	3,057.1	(10,24)	2,139.9
2	(5,16)	3,048.3	(13,24)	2,241.8
3	(6,9)	3,038.4	(10,13)	2,307.1
4	(2,6)	3,030.6	(24,32)	2,380.2
5	(8,35)	3,026.3	(10,15)	2,380.3
6	(2,8)	3,024.6	(10,11)	2,380.6
7	(16,38)	3,022.4	(10,32)	2,401.3
8	(6,34)	3,013.3	(1,10)	2,411.4
9	(6,17)	3,007.8	(10,30)	2,412.0
10	(2,5)	3,005.8	(22,23)	2,440.7

(2) Results for six months and 12 months

Table 4.10 shows the EES with combination of nodes in a 6-month period. The combination of nodes 8 and 35 becomes the top one, leading to the EES of $7,137.4$ kWh which is about 42% larger than the worst combination of nodes 13 and 24. Compared to the best single information source scenario for 6-month, as shown in Table 4.5, the best combination of two source nodes achieve 4.75% more EES.

Table 4.10: EES of Top and Bottom 10 Pairs of Information Sources within 6 Months

Rank	Top Pairs	EES (kWh)	Bottom Pairs	EES (kWh)
1	(8,35)	7,137.4	(13,24)	5,034.3
2	(6,8)	7,053.4	(17,24)	5,508.4
3	(8,26)	7,037.2	(23,24)	5,545.3
4	(2,8)	7,027.0	(10,25)	5,601.4
5	(16,38)	6,999.9	(13,17)	5,646.4
6	(8,31)	6,992.6	(15,25)	5,677.0
7	(2,16)	6,975.6	(25,30)	5,699.1
8	(5,16)	6,964.2	(24,25)	5,703.4
9	(8,28)	6,941.7	(10,13)	5,717.0
10	(9,26)	6,938.7	(1,10)	5,731.4

Table 4.11: EES of Top and Bottom 10 Pairs of Information Sources within 12 Months

Rank	Top Pairs	EES (kWh)	Bottom Pairs	EES (kWh)
1	(8,35)	15,808	(13,24)	12,136
2	(6,8)	15,713	(16,24)	12,241
3	(8,26)	15,708	(17,24)	12,411
4	(16,38)	15,688	(16,17)	12,477
5	(8,31)	15,657	(13,17)	13,004
6	(2,8)	15,651	(13,16)	13,234
7	(8,38)	15,601	(10,24)	13,421
8	(8,28)	15,581	(10,13)	13,796
9	(8,12)	15,573	(10,17)	13,957
10	(8,39)	15,508	(1,24)	14,006

Additionally, the normalized difference of EES between the top and bottom node pairs is decreasing when the time period is longer. This indicates that the social network with two information sources is likely to become saturated shortly after 3 months. Table 4.11 shows that the best combination of nodes 8 and 35 can save $15,808$ kWh in a whole year, which is 30% more than the worst combination of nodes 13 and 24.

(3) EES with best and the randomly selected group of source node pairs

Table 4.12 is a comparison of the EES between the best combination and the average of 10 randomly selected combinations. It can be seen that in the first month, the best combination has 47.40% more savings than the randomly selected group. At the end of one year, it is still 6.77% higher than the random group. This implies the benefits of choosing the best combination of source pairs in achieving targeted energy savings.

Table 4.12: Comparison of EES with Best Combinations to the Average of 10 Random Combinations

Months	Best Pair (kWh)	Ave. of 10 Random Pairs (kWh)	Improvement
1	691.2	468.9	47.40%
2	1,833.8	1,567.9	16.96%
3	3,057.1	2,722.7	12.28%
6	7,137.4	6,594.5	8.23%
12	15,808.0	14,806.0	6.77%

4.4.6 Cost Benefit Analysis

A cost saving calculation is derived for decision makers managing a mass roll-out programme. Define the electricity tariff to be constant at £π per kWh, and the cost of one piece of free trial product to the roll-out organiser is C_0, each source node will adopt l_p pieces of product. The total EES can be calculated as $\pi \tilde{F}^T$, the accumulation from $t = 1$ to $t = T$, and the total cost benefit can be calculated as

$$C(T) = \pi \cdot \tilde{F}^T - C_0 \cdot l_p \cdot \|\Xi\|. \tag{4.23}$$

The electricity tariff £π is the same as given in Section 4.4.1, which is electricity tariff $\pi = 0.16$ £/kWh. There are 12 LED bulbs and each one costs £4.97, that is, $l_p = 12, C_0 =$£4.97. When calculating the cost benefit using (4.23), the values of EES can be found from the result tables in Section 4.4.3.2 for single source node and Section 4.4.5 for double source nodes.

Taking node 6 as the single information source node, the cost benefit achieved for 1 month is calculated to be £43.49, which means that when taking node 6 as the information source node, the payback period is less than one month as the value is larger than 0 (the cost of EES in currency larger than the initial cost). Keeping node 6 as the source node, the cost benefit achieved for three months is calculated to be

£411.99. The cost benefit of taking node 5 as the source node is calculated to be £1,030.52 for 6 months, and the cost benefits for node 8 is calculated to be £2,424.52 for 12 months. The above details for single information source node and the following about two source nodes can be seen clearly in Table 4.13.

For the analysis with two source nodes, the cost benefit for the best and worst combinations are -£8.70 and -£82.43, respectively, which indicates that the payback period of dual information sources is longer than one month. The reason for having negative cost benefit is that the initial cost of two source nodes is twice as much as the single node, while the savings from EES one month is not enough to cover the cost spent on EEP within this short period of time. When the running time is longer, more savings will be obtained to compensate the investment on EEP of two source nodes, therefore getting positive benefits.

Table 4.13: Cost Benefit Analysis of Single and Dual Information Sources

Months	Best Single or Pair of Nodes	EES (kWh)	Cost Benefit (£)
1	(node 6)	644.6	43.5
	(16,38)	691.2	-8.7
3	(node 6)	2,947.7	412.0
	(6,8)	3,057.1	369.9
6	(node 5)	6,813.5	1,030.5
	(8,35)	7,137.4	1,022.7
12	(node 8)	15,526.0	2,424.5
	(8,35)	15,808.0	2,410.0

The cost benefits for the best and worst combinations for three months are £369.86 and £223.10, respectively. For six months, the cost benefits for the best and worst combinations are £1,022.70 and £686.21, respectively. For 12 months, the cost benefits for the best and worst combinations are £2,410.00 and £1,822.48, respectively. It is apparent that the cost benefit is larger when the running period is longer.

When comparing single source and dual source information nodes, it becomes evident that a single source offers advantages in cost benefit compared to utilizing two information sources for a small social network. When the network is larger, it may need more source nodes to obtain the best cost benefits. This can be further investigated following an optimization as suggested in Section 4.3.

4.5 Summary

A new network model to calculate the EES is proposed in this chapter by studying user interactions within a social network. The dynamic propagation of information within the social network can be described. With the use of probability theory, the established mathematical model can be used to calculate both the direct and indirect energy savings over a given time period, for both single and multiple information sources. This model can also help to decide which individuals or combined individuals will have the most influence on the social network. The model is tested by a social network with 40 participants and the network data is collected from survey questions.

When a single information source is considered, the case study results show that the influence of the source node is not only related to its number of connections, but is also strongly affected by node to node connection strength, measured by weighting coefficients. Using the identified best information source, the energy savings in two months is about 25% more than randomly chosen nodes on average. At the end of a one-year period, it still can have 5.64% more energy savings than randomly chosen nodes on average. Therefore, the best node identified by the model can achieve significant amount of extra energy savings in a short time period.

When the network has multiple information sources, the influence of these multiple source nodes depends on the connections and weights of the network, and the combination of the two most influential single information source nodes may not provide the greatest influence to the network. Within one month, the best combination of information sources can achieve about 47% more energy savings compared to randomly chosen group of combinations on average. This again illustrates the benefits of using the model-based analysis results in guiding energy efficiency technology roll-out projects. The model will help to monitor the progress of the installations of EEP, and analyze if the targeted progress can be achieved. The obtained model can be further expanded and implemented for large-scale energy related programmes and projects, for example, the roll-out of smart meters, electric vehicles, etc.

The modelling method presented in this chapter has a few limitations. Firstly, the method requires survey data including details of social network connections. This narrows the scale of the targeted social network as such data are not always available for large networks. In the next chapter, epidemic theory will be applied to minimise the requirement of network connection details. Sampling techniques from statistics can also be employed to reduce the number of man power involved in a survey. Secondly, there are a large number of parameters about the network to be determined in the presented model, which brings computational challenges and increase the model uncertainty when the data set is small for parameter identification. The current approach in the case studies of this chapter is to select sufficiently large populations in the generic algorithm to solve the least square optimization problem. More advanced optimization algorithms need to be developed if the size of the network is very big. Thirdly, the reliability of the EES depends on whether the sampled social network is representative enough. Inappropriately selected network may cause the results less representative. Advanced sampling techniques in (Cochran, 1977; Ye et al., 2014) can be applied to guide the selection of individuals and the corresponding network.

Chapter 5

Investigating Impact of Personal and Social Network Factors to Adoption of Energy Products

5.1 Introduction

Moving forward from the modelling and analysis of energy savings in a social network, the main research focus in this chapter is to investigate the impact of individual user factors as well as social networks influence on on energy savings.

The impact of user behaviours has drawn recent research attention in the area of energy innovation. Each individual's decision plays a key role towards improving energy use efficiency. Power companies and engineers are keen to understand how user behaviour may shape energy consumption and how to influence users to reduce energy consumption from social sciences point of view (Bavaresco et al., 2020). To achieve effective energy savings, it is imperative to conduct research on both social and technical scales (Figueiredo et al., 2005; Fischer, 2008). The social scale would detail how consumer behaviour in energy conservation would be affected by information sharing within the social network (Boyd and Ellison, 2007). The technical scale can recommend how such social networks and information shared can be modelled (Ellison et al., 2007; Feng et al., 2013).

It is understood that personal acceptance and social network influence will change householder perception and consumer behaviour towards energy savings. The personal acceptance level is related to income, age, family situation and more. While user behaviour affects choices when it comes to energy conservation, information to control or improve energy saving behaviour is a significant parameter that is difficult to obtain. There's a lack of data and models to represent the influence of consumer behaviours on the energy savings outcome. A quantitative research method is particularly useful when the objective is clearly stated, and a hypothesis-based testing is required.

In this chapter, a mathematical model will be developed that can be used to quantify the influence of household characteristics such as income, age, education, etc., the social network influence from nearby and wider network, and possible external influence on energy savings. Model-based analysis and design results can be used to guide innovation program in residential energy savings. A big challenge to this quantitative methodology is to understand the requirement of model complexity and collect real-world data to support model development.

The remainder of the Chapter is structured as follows. A mathematical model is developed in Section 5.2 to describe the dynamics of EEP adoption in social networks considering various individual factors, social network influence and advertisement impact. Case studies including a small-scale network, the same one as used in Chapter 4, and a large population network, established through a new survey, are conducted in Section 5.3, with results and new insights provided. A summary is given in Section 5.4.

5.2 Methodology on Model Development

5.2.1 Network Modelling

A social network with N householders can be represented as a network with N nodes, the connections between householders are represented by edges. The connection state of the whole network can be represented by a $N \times N$ matrix $\mathbf{A}$, in which A_{ij} denotes the edge state from node i to node j,

$$A_{ij} = \begin{cases} 1, & \text{when node } i \text{ is connected to node } j \\ 0, & \text{otherwise} \end{cases} \tag{5.1}$$

In this chapter, the influence strength between different nodes are taken to be the same. Thus, no weight coefficients are added to edges, A_{ij} is therefore a binary term. This is different from the a_{ij} used in Chapter 4. Another assumption is that the connections between a pair of nodes are bidirectional, that is, $A_{ij} = A_{ji}$, which follows connections for general networks. The degree of each node (number of nodes connected), k_i, can be calculated from A_{ij} as,

$$k_i = \sum_{j=1}^{N} A_{ij}. \tag{5.2}$$

State variable $n_i(t), (i = 1, 2, \ldots, N)$, are defined to describe the householder status on adoption of EEP at each time.

$$n_i(t) = \begin{cases} 1, & \text{when node } i \text{ has adopted EEP;} \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

Here $n_i(t)$ is the state of node i at time t.

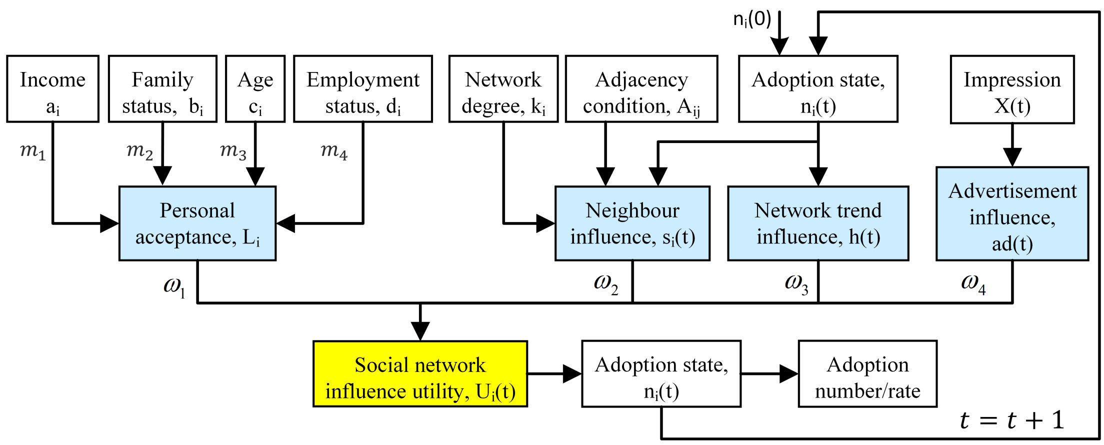


Figure 5.1: Model Configuration

In this work, both individual influence and social network influence are considered for adopting EEP. Four factors are included in the modelling and analysis, they are personal acceptance level, influence from neighbours, influence from social network trend, and advertisement influence. The model configuration is shown in Fig. 5.1. More details on each of the factor will be presented in the following sections.

5.2.2 Personal Acceptance Level Model

When considering the individual user factors that affect EES, four factors are selected in this study, which are income level, family status (having children or not), age group, and employment status. A simple linear function is defined here with the personal acceptance level denoted by L_i for node i. To enable the modelling, the four local factors influencing personal acceptance level are classified as follow (NOP and Research, 2012).

- Income. The state of income level is denoted by $a = 1, 2, 3$, representing household income being low (below £17K pa), mid (between £17K pa to £37K pa), and high (above £37K pa), respectively.
- Family status. The state of family situation is described by $b = 1, 2$, representing household having no children and having children, respectively.
- Age group. The age group is classified into four groups, represented by $c = 1, 2, 3, 4$, for age between 18 to 34, 35 to 54, 55 to 64, and above 65.
- Employment status. The state of working status is represented by $d = 1, 2$ for unemployed and employed, respectively.

The contribution of each local factor to the personal acceptance level is associated by a weighting factor, and linearly integrated together as follows:

$$L_i = m_1 a_i + m_2 b_i + m_3 c_i + m_4 d_i, \tag{5.4}$$

where m_1, m_2, m_3, and m_4 are weighting coefficients for income, family situation, age and employment status, respectively, taking values between -1 to 1.

The personal acceptance level could be difference from one node to another, they are not changed with time. The sub-system model on personal acceptance level is therefore a static model. It is assumed that the same set of parameters apply to all nodes. In this proposed model (5.4), the unknown parameters (m_1, m_2, m_3, m_4) are estimated using information abstracted from the survey data (NOP and Research, 2012). Denote

$$\boldsymbol{\theta}_L = [m_1, m_2, m_3, m_4]^\top . \tag{5.5}$$

The parameters in $\boldsymbol{\theta}_L$ can be determined through an LS estimation problem, i.e.,

$$\begin{aligned} &\boldsymbol{\theta}_L^* = \arg\min_{\boldsymbol{\theta}_L \in \boldsymbol{\Theta}_L} \textstyle\sum_{t=1}^{N_L} \left(m_1 \overline{a}(t) + m_2 \overline{b}(t) + m_3 \overline{c}(t) + m_4 \overline{d}(t) - \tilde{y}(t) \right)^2 \\ &\text{s.t.} \quad m_1 \overline{a}(t) + m_2 \overline{b}(t) + m_3 \overline{c}(t) + m_4 \overline{d}(t) \leq 1 \end{aligned} \tag{5.6}$$

where t is the index number for the data sequence, N_L is the total number of data for modelling, $\overline{a}(t)$, $\overline{b}(t)$, $\overline{c}(t)$, and $\overline{d}(t)$ are the averaged values of four factors at each k, $\boldsymbol{\Theta}_L$ is the searching domain for $\boldsymbol{\theta}_L$, $\tilde{y}(t)$ are acceptance levels provided in the surveyed data.

5.2.3 Neighbouring Network Influence Model

The adoption state of a node in the network can be represented as a binary variable $n_i = 0, 1$, where 1 and 0 representing adopted and not adopted. Initially, some nodes in the network have already adopted the EEP, they are taken as the seed nodes. These seed nodes will share EES information to neighbour nodes, which may influence the state of their neighbours. When the states of neighbour nodes become to 1 from 0, triggered when the adoption condition is satisfied, they will become the new seed nodes and exert influence on their neighbour nodes.

The initial state of seed nodes are randomly distributed among the network. The initial seed nodes take a certain proportion of the nodes in the whole network. At each time t, the influence of neighbouring adoption to node i is quantified as a ratio between the number of adopted neighbours to the node degree of i, which is characterized by

$$s_i(t) = \frac{1}{k_i} \sum_{j=1}^{N} A_{ij} n_j(t), \ (j \neq i). \tag{5.7}$$

The neighbour influence model (5.7) is a dynamic model, the change of s_i(t) depends on the update of $n_i(t)$ at each time t.

5.2.4 Network Trend Model - Network Adoption Rate

It is indicated in (Bale et al., 2014) that the entire social network has influence on householder's decision making. This majority affect phenomenon can be understood as the householder in a social network has desire to adapt to the society (Bale et al., 2014). It can be taken as indirect influence from the whole network apart from the direct influence from neighbouring nodes. This indirect influence is named as network trend influence, denoted by $h(t)$ in this chapter, where it is quantified as the same value as the EEP adoption rate, denoted by $R(t)$, of the network at time t,

$$h(t) = R(t) = \frac{1}{N} \sum_{i=1}^{N} n_i(t) \tag{5.8}$$

Since the node state $n_i(t)$ changes over time, the network trend influence $h(t)$ is also a function of time.

5.2.5 Advertisement Influence Model

The use of advertisement shows significant influence to the EEP adoption rate (Du et al., 2023; Shapiro, 1983). Advertisement input is considered as an external, manageable factor in this thesis work.

The 'exponential family model' is employed to establish the advertisement model. The exponential family model is essentially a broader, more generalized dynamic linear model within the Bayesian forecasting theory BFT framework. An exemplary implementation of such a model was provided by West and Harrison (1997), where they elucidated a case study centered on TV advertisement awareness. Building upon this model, modifications were introduced to forecast online advertisement effectiveness.

Denote the number of advertisement views at time t as $X(t)$, the adoption rate as $\mu(t)$. Here, the immediate assumption is that the present $X(t)$ will directly escalate the advertisement's influence on product adoption. To compute the advertisement effectiveness, certain parameters and settings are deemed necessary.

- Lower threshold, α

  Represented by parameter α, the lower threshold encapsulates the minimum possible advertisement adoption rate at any time, such that $0 \leq \alpha \leq 1$. In this PhD study, the lower threshold's initial value is posited to be the product's starting adoption rate.

- Upper threshold, β

  The upper bound for the advertisement adoption rate is denoted by parameter β, ensuring $0 \leq \mu(t) \leq \beta \leq 1$. Based on the guidelines in West and Harrison (1997), the value of β is set to be 0.85 in this thesis work.

- Decay factor, ρ

  Acknowledging the inevitable memory decay associated with innovative information, it's assumed that the adoption rate undergoes a decay between any two consecutive times. This is characterized by a factor of $100(1-\rho)\%$, where ρ is constrained between $0 \leq \rho \leq 1$, and represents the decay factor for $\mu(t)$. In alignment with (West and Harrison, 1997), the value for ρ is set at 0.90.

- Diminishing return factor, γ

  Echoing the economic concept of diminishing returns, it's posited that if one element of production escalates while others remain static, the resulting benefits will diminish gradually. In relation to this model, it implies that as impression $X(t)$ increases, the advertising effect is amplified by a fraction of the residual awareness. This fraction is gauged by the penetration function $\exp(-\gamma X(t))$, which exponentially decays in tandem with the rise of $X(t)$, where $\gamma \geq 0$. In this model, γ is coined as the diminishing return factor and is initialized at 0.02, as referenced in West and Harrison (1997).

Given the above four parameters, the effect of advertisement impression at time t, $E(t)$, can be developed using the BFT as

$$E(t) = (\beta - \alpha) - (\beta - \alpha - \rho E(t-1)) \cdot \exp\left(-\gamma X(t)\right). \tag{5.9}$$

The advertisement influence, $ad(t)$, is calculated as the sum of lower threshold, α, and the effect of impressions, $E(t)$, which is shown in (5.9), i.e.,

$$\begin{aligned} ad(t) &= \alpha + E(t) \\ &= \beta - \Big(\beta - \alpha - \rho E(t-1) \cdot \exp\big(-\gamma X(t)\big)\Big) \end{aligned} \tag{5.10}$$

where α and β are lower and upper thresholds, ρ is a decay factor of advertisement effect, γ is a diminishing return factor of advertisement effect. The values of α, β, ρ are between 0 and 1, while γ is larger than 0.

Model (5.10) is the advertisement model, describing the relationship between the impression $X(t)$ (number of views, or equivalently advertisement input) and the effect of advertisement. This advertisement model is not affected by the network system, its dynamics is determined by the model parameters $(\alpha, \beta, \rho, \gamma)$. The initial value of $E(t)$ is assumed to be the lower threshold obtained in (Du et al., 2023) as no nodes receives advertisement at the initial stage, where $E(0)$ is $1.697e-2$ for the small network, and $3.815e-6$ for the large network.

5.2.6 Adoption Utility

In order to examine the combined contribution from personal influence level, neighbouring network influence, network trend influence and advertisement influence to the EEP adoption decision making of every node in a network, the following utility measure is defined.

$$U_i(t) = \omega_1 \cdot L_i + \omega_2 \cdot s_i(t) + \omega_3 \cdot h(t) + \omega_4 \cdot ad(t) \tag{5.11}$$

where ω_1, ω_2, ω_3, and ω_4 are the relative weights of personal acceptance level, neighbouring network influence, network trend influence and advertisement influence, and $\omega_1 + \omega_2 + \omega_3 + \omega_4 = 1$.

5.2.7 Decision of Adoption - Update of Node State

When the adoption utility reaches the defined threshold ε at time t, node i will adopt the EEP which results in the change of n_i from 0 to 1 at time $t+1$. As the decision of adoption is a process of information accumulation, the state dynamics of node i can be represented as,

$$n_i(t+1) = \begin{cases} 1, & \text{if } U_i(t) \geq \varepsilon \\ 0, & \text{if } U_i(t) < \varepsilon \end{cases} \tag{5.12}$$

The change of state n_i will be updated in the neighbour influence model (5.7) and the network trend model (5.8), the evolution will continue until the system reaches steady state.

The model configuration is shown in Fig. 5.1. The algorithm flowchart to calculate the decision of adoption from time t to $t+1$ is shown in Fig. 5.2, and the pseudo-code is given in Algorithm 5.1.

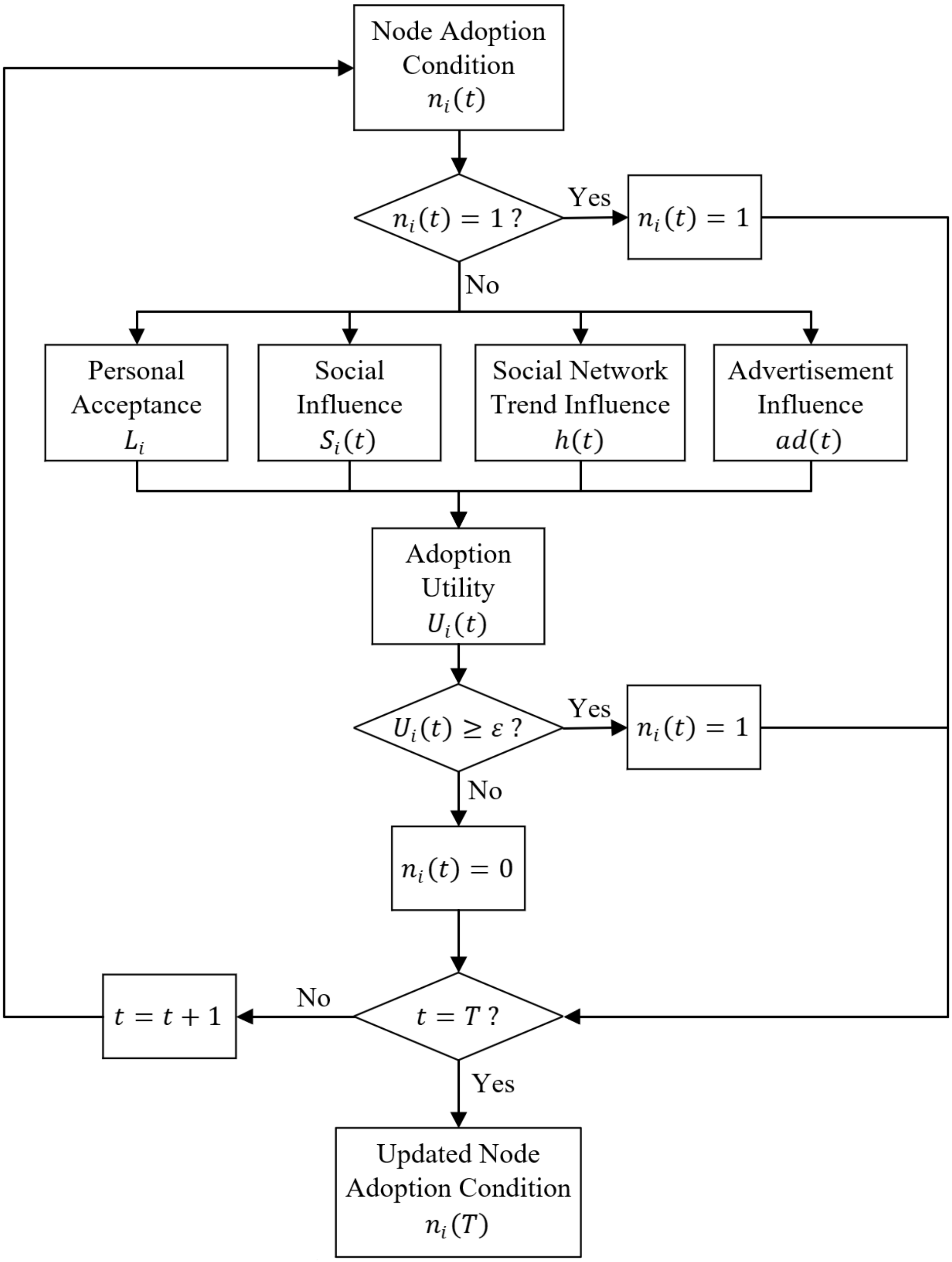


Figure 5.2: Decision of Adoption Algorithm Flow Chart

Algorithm 5.1 Updating Adoption Rate by Using Adoption Utility

Initialise network connection condition A_{ij}
while time=t **do**
  Update current adoption rate $h(t)$, see equation (5.8)
  Update current Advertisement influence $ad(t)$, see equation (5.10)
  for each source node i **do**
    Calculate personal acceptance L_i for node i, see equation (5.4)
    for Each neighbour node j **do**
      Calculate influence of neighbouring node $s_i(t)$, see equation (5.7)
    end for
  end for
  Calculate adoption utility $U_i(t)$, see equation (5.11)
  Check if $U_i(t) \geq \varepsilon$?
  If yes, update adoption rate $n_i(t+1) = 1$
  Otherwise, update adoption rate $n_i(t+1) = 0$
end while
$t = t + 1$, move to the next time iteration until the simulation is completed

5.2.8 Determination of Unknown Model Parameters

In order to simulate the adoption behaviour using the proposed model, the unknown parameters in the advertisement model, $(\alpha, \beta, \rho, \gamma)$, need to be determined. We also need to set the weighting coefficients in the utility function (5.11), $(\omega_1, \omega_2, \omega_3, \omega_4)$, so that the value of $U_i(t)$ can be calculated. Another parameter to set up is the decision making threshold ε in the node state update equation, (5.12). In this work, the values of these parameters are determined by LS parameter estimation using the survey data.

Denote

$$\boldsymbol{\theta}_U = [\varepsilon, \alpha, \beta, \rho, \gamma, \omega_1, \omega_2, \omega_3, \omega_4]^\top . \tag{5.13}$$

$\boldsymbol{\theta}_U$ is the vector of unknown parameters for the proposed model, which can be estimated by solving the following optimisation problem,

$$\boldsymbol{\theta}_U^* = \arg \min_{\boldsymbol{\theta}_U \in Z_{\boldsymbol{\theta}_U}} J_{\boldsymbol{\theta}_U}(\boldsymbol{\theta}_U) \tag{5.14}$$

$J_{\boldsymbol{\theta}_U}$ is the objective function to be minimised , the set $Z_{\boldsymbol{\theta}_U}$ is the searching domain for $\boldsymbol{\theta}_U$, which is a set of positive real numbers in the range between 0 to 1. The objective function is formulated as the residual function in LS estimation, that is,

$$
\begin{aligned}
J_{\boldsymbol{\theta}_U} &= \textstyle\sum_{t=1}^{T} \left(R(t,\boldsymbol{\theta}_U) - y_R(t)\right)^2 \\
&\text{subject to } \omega_1 + \omega_2 + \omega_3 + \omega_4 = 1
\end{aligned}
\tag{5.15}
$$

where $y_R(t)$ is the adoption rate data obtained from survey at time t, $R(t)$ is the adoption rate data calculated by the model at time t, T is the total number of time points. The algorithm used for parameter estimation is shown in Algorithm 5.2.

Algorithm 5.2 GA Algorithm for Social Network Influence Utility Parameter Estimation

```
1: Initialisation
2: t ← 0
3: Generate population P_t with 50 feasible solutions randomly
4: while J_θU(θ_U)>1e − 6 do
5:     Select solutions from P_t
6:     Recombine solutions
7:     Mutate solutions
8:     P_{t+1} ←newly created solutions
9:     t ← t + 1
10: end while
11: Output best solutions in P_t
```

5.3 Case Study with Two Networks

5.3.1 System Setup

In this chapter, case studies on a small and a large population networks are analysed. Three survey data sets are used.

- The small population network used in this case study is based on a survey of 40 participants (Du et al., 2016). This is the same social network as surveyed in Chapter 4, but more data are used in this chapter. In addition to the user connections, user response to recommendation of a new EEP through the network influence, their response to advertisement exposure frequencies are surveyed.

- The large population network used is based on one million users in a large trading platform. The collected user data do not contain connection details, therefore a complex network modelling method is applied to establish the network structure.

- In order to model the personal acceptance level considering individual (local) factors, the survey data of Green Deal Segmentation(GDS) (NOP and Research, 2012) was used.

5.3.2 Personal Acceptance Level

In the research report of GDS (NOP and Research, 2012), survey data are collected for the willingness of new EEP adoption. Participants' income level, age, family situation and working status are collected as well. In order to find out the relative contribution of these householder circumstances to the acceptance rate of EEP, the following assumptions are made.

1. All the nodes take the same values of weighting parameters for (m_1, m_2, m_3, m_4) in L_i, (5.4).

2. The personal acceptance in the report of GDS is divided into six different levels with relative weights of 0, 0.2, 0.4, 0.6, 0.8 and 1, respectively, which represents rejecters, low acceptance, below average acceptance, above average acceptance, high acceptance, and accepters. The proportions of each segments of acceptance level can be seen in Table 5.1, from which it is observed that 11% users reject to adopt EEP regardless of any influence, 20% users will definitely adopt EEP, other people are in between, their decision will be affected by network and advertisement influence.

3. For simplification, averaged values of four local factors, i.e., income level, age, family situation, and working status, are used for each of the six levels of personal acceptance. These mean values are used in modelling of the personal acceptance level subject to the four local factors.

Table 5.1: Acceptance Levels of All Segments

Segments Index	1	2	3	4	5	6
Acceptance Level, L	0	0.2	0.4	0.6	0.8	1
Proportion in Population (%)	11	24	11	10	24	20

The proportion of each factor in each segment are collected and listed in Table 5.2.

Table 5.2: Proportion of Personal Factors (%) and Average Values

Personal Factors	1	2	3	4	5	6
Low Income ($a = 1$)	38	41	30	11	23	32
Mid Income ($a = 2$)	30	40	46	30	39	42
High Income ($a = 3$)	32	19	24	59	38	26
Average ($\bar{a}$)	1.94	1.78	1.94	2.48	2.15	1.94
No Children ($b = 1$)	62	76	68	50	74	43
Have Children ($b = 2$)	38	24	32	50	26	57
Average ($\bar{b}$)	1.38	1.24	1.32	1.50	1.26	1.57
Age group 1 ($c = 1$)	17	23	30	36	35	38
Age group 2 ($c = 2$)	46	24	36	50	30	44
Age group 3 ($c = 3$)	12	11	12	10	19	10
Age group 4 ($c = 4$)	25	42	22	4	16	8
Average ($\bar{c}$)	2.45	2.72	2.26	1.82	2.16	1.88
Unemployed ($d = 1$)	45	54	38	19	32	30
Employed ($d = 2$)	55	46	62	81	68	70
Average ($\bar{d}$)	1.55	1.46	1.62	1.81	1.68	1.70

The values in $\boldsymbol{\theta}_L$ can be obtained by the LS identification algorithm in (5.6) using the processed data in Table 5.2. In the Matlab simulation, the 'fmincon' function is chosen to find the solution leading to the minimum objective function. The initial values of $\boldsymbol{\theta}_L$ used for 'fmincon' is the result from a GA optimization. The evolution profiles of parameter estimation are shown in Fig. 5.3, from which it can be seen that m_1 and m_3 achieved the steady states after 80 iterations while m_2 and m_4 after 95 iterations.

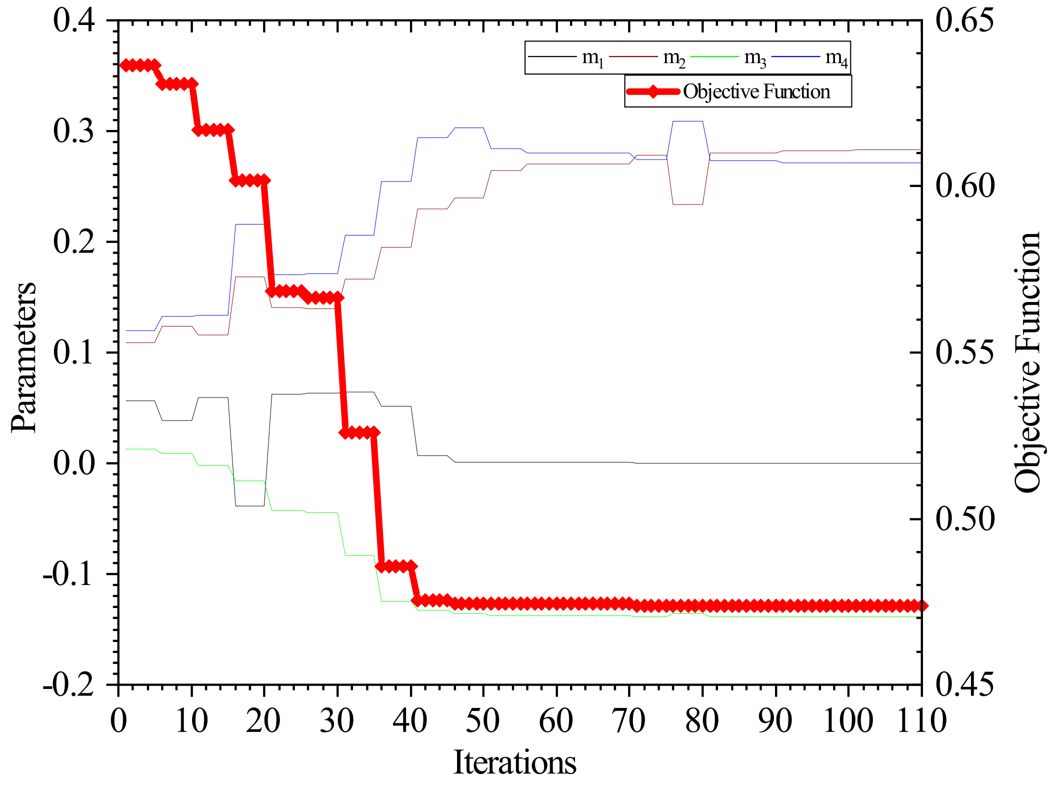


Figure 5.3: Convergence Graph along Iterations in Fmincon Function of Parameter Estimation (m_1, m_2, m_3, m_4) Using the Adoption Rate Model

The estimation results for (m_1, m_2, m_3, m_4) are shown in Table 5.3. It can be seen that, compared to the other three factors, the coefficient of income m_1 has a very low magnitude, indicating that the personal acceptance level is not much dependent on income. User families with children have more interest in adopting EEP (m_2), so do those with a job (m_4). The coefficient m_3 has a negative value, implying that younger people are more likely to adopt EEP.

Table 5.3: Coefficients of Segments

Local factors	Income	Family	Age	Employment
Coefficients	m_1	m_2	m_3	m_4
Initial values	$4.283e-6$	0.1347	-0.1287	0.3802
Final values	$4.671e-7$	0.2830	-0.1385	0.2710

Once the four coefficients (m_1, m_2, m_3, m_4) are determined and the (a, b, c, d) data for each individual factor are given, the personal acceptance level, L_i, can be calculated for each node in the social network with the model in (5.4).

5.3.3 Advertisement Input Setup

The advertisement input (impression, or number of views), $X(t)$, is an external input factor required to calculate the advertisement effect $E(t)$ in (5.9), and eventually the advertisement influence $ad(t)$ in (5.9). For the small network, $X(t)$ is set as one day per week with the same magnitude for each input (Du et al., 2016). For the large social network, the advertisement input data is obtained from the survey data and given in Appendix A.

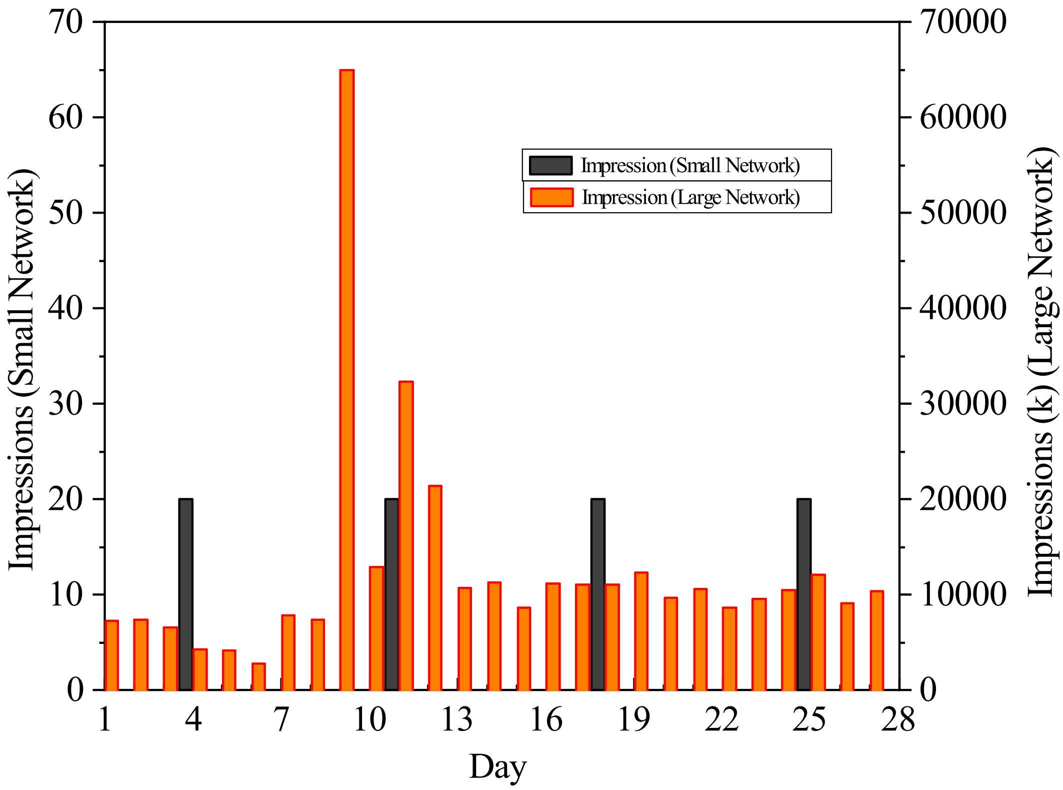


Figure 5.4: Impressions for Small and Large Population Network

The input profiles of $X(t)$ for the two networks are shown in Fig. 5.4 for a period of 28 days. It can be seen that in a large population network, the advertisement impressions are consistently invested from day 1 to day 28. They are used as baselines for advertisement input.

5.3.4 Small Population Network

5.3.4.1 Baseline System Simulation

In this survey, all 40 participants are PhD students thus the values of the four local factors are the same for personal acceptance level: $a = 1$ (low income), $b = 1$ (no children), $c = 1$ (age 18-34), and $d = 1$ (unemployed).

There is limitation on the data of small population network, as the participants are a group of young adults (PhD students). However, the following cases are used to compare the effectiveness of advertisement. The results can still yield a satisfactory outcome. The parameters of the personal acceptance level (m_1, m_2, m_3, m_4) are those identified values listed in the last row of Table 5.3.

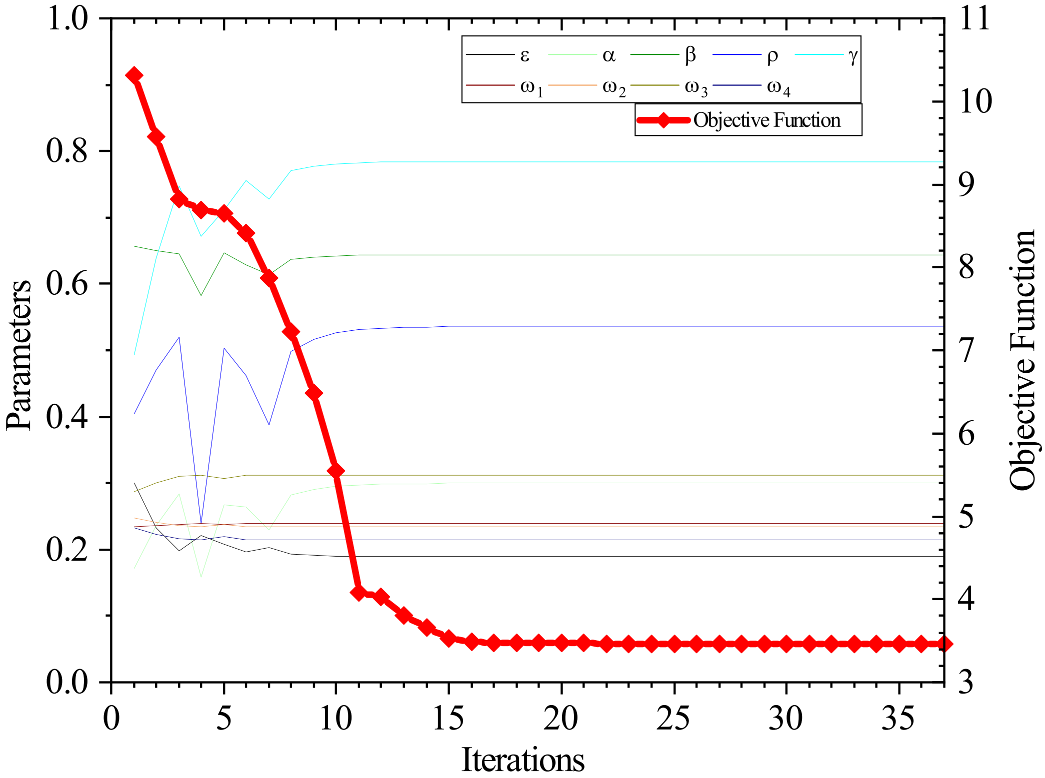


Figure 5.5: Parameter Estimation along Iterations $(\varepsilon, \alpha, \beta, \rho, \gamma, \omega_1, \omega_2, \omega_3, \omega_4)$ for Small Population Network

The unknown parameters in $\boldsymbol{\theta}_U$ are obtained using the LS estimation in (5.14) by applying data from the survey. The evolution profiles of parameter estimation is shown in Fig. 5.5. It can be seen that all nine parameters reach the steady states after 17 iterations. The 'fmincon' is applied in Matlab simulation.

Table 5.4: Small Network Model Parameters Estimation Result

ε	α	β	ρ	γ
0.1891	0.2998	0.6440	0.5358	0.7843
ω_1	ω_2	ω_3	ω_4	
0.2388	0.2344	0.3127	0.2141	

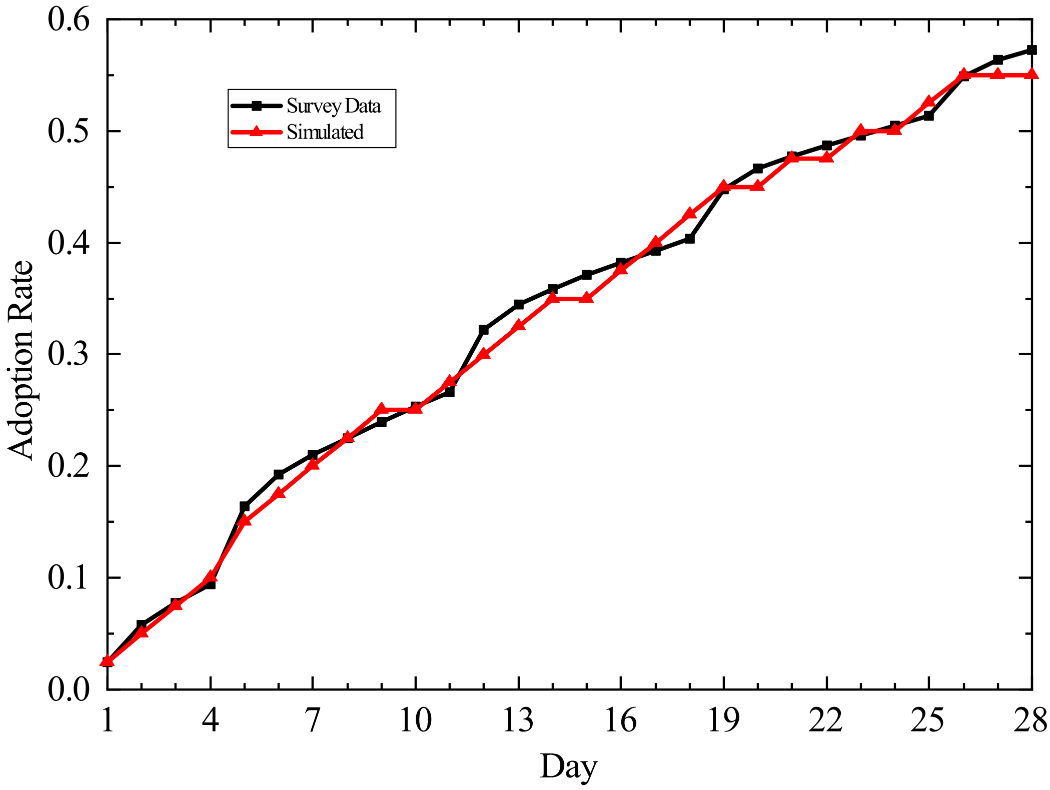


Figure 5.6: Small Network Adoption Rate: Comparison between Survey Data and Model Outcome

The optimised results for parameter estimation are shown in Table 5.4, from which it can be seen that the four weighting coefficients for the social network influence utility, $(\omega_1, \omega_2, \omega_3, \omega_4)$, have values between 0.2141 to 0.3127, among them ω_3 is the largest indicating a larger influence of the network trend on the adoption rate. The threshold ε for small population network is found to be 0.1891 from the estimation.

Using these identified parameter values, the simulated adoption rate is calculated from the proposed model, and the results are compared with the adoption rate data in survey (see Fig. 5.6). A good matching can be observed between the modelling results and the survey data.

5.3.4.2 Advertisement Control Practice

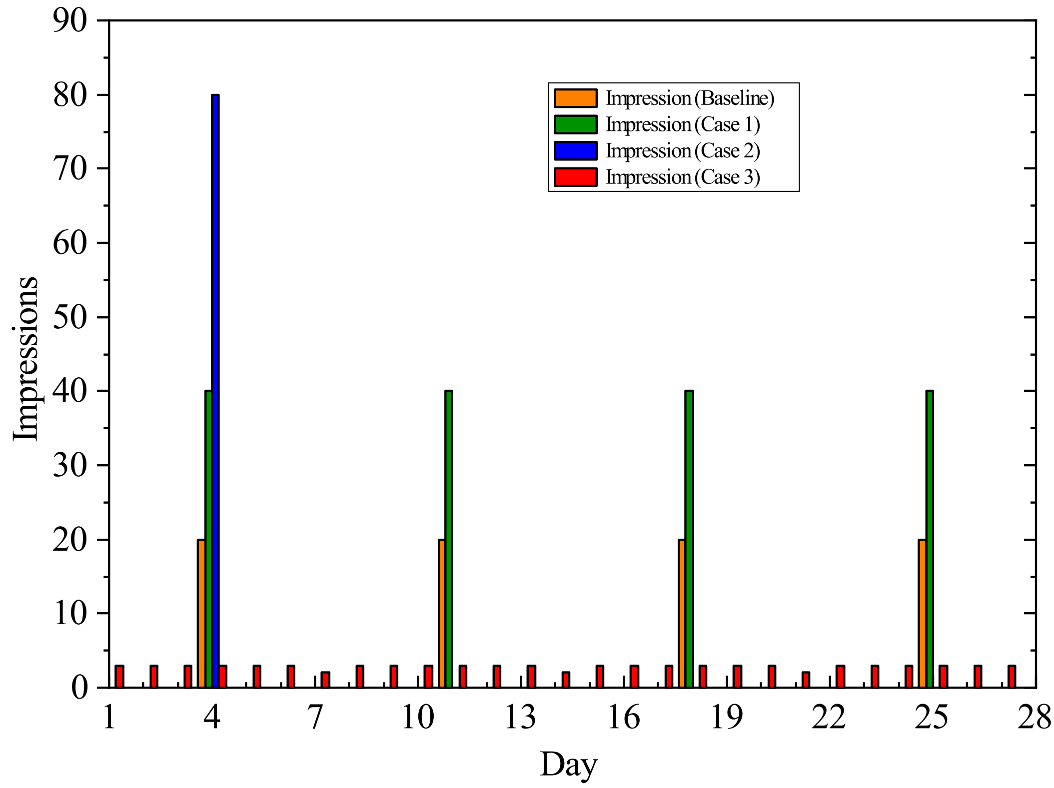


Figure 5.7: Four Profiles of Advertisement Input on Small Network

After obtaining the weighting coefficients for the social network influence utility, simulations with different advertisement inputs, i.e., impression $X(t)$, are conducted for the small network. It includes one baseline input and three other input profiles.

- Baseline. The advertisement input is to applied once per week with the magnitude of 20 for each entry, as shown in in Fig. 5.7 with bars in orange colour.
- Case 1. The advertisement input frequency is the same as the baseline, which is once per week, but the impression magnitude is twice of the baseline magnitude, as indicated by the green bar in Fig. 5.7.
- Case 2. There's a single advertisement input applied in the first week, the impression magnitude is 80, which is equivalent to the 4 entries putting together in the baseline profile. The Case 2 input is represented by the blue bar in Fig. 5.7.

- Case 3. The total amount of 80 impression in the baseline are split into daily actions, which is shown as red bars in Fig. 5.7.

A detailed breakdown of the daily impression distribution for the baseline and the three cases can be found in Table 5.5.

Table 5.5: Daily Impression of Advertisement Practice on Small Network

Day	1	2	3	4	5	6	7
Impressions (Baseline)	0	0	0	20	0	0	0
Impressions (Case1)	0	0	0	40	0	0	0
Impressions (Case2)	0	0	0	80	0	0	0
Impressions (Case3)	3	3	3	3	3	3	2
Day	8	9	10	11	12	13	14
Impressions (Baseline)	0	0	0	20	0	0	0
Impressions (Case1)	0	0	0	40	0	0	0
Impressions (Case2)	0	0	0	0	0	0	0
Impressions (Case3)	3	3	3	3	3	3	2
Day	15	16	17	18	19	20	21
Impressions (Baseline)	0	0	0	20	0	0	0
Impressions (Case1)	0	0	0	40	0	0	0
Impressions (Case2)	0	0	0	0	0	0	0
Impressions (Case3)	3	3	3	3	3	3	2
Day	22	23	24	25	26	27	28
Impressions (Baseline)	0	0	0	20	0	0	0
Impressions (Case1)	0	0	0	40	0	0	0
Impressions (Case2)	0	0	0	0	0	0	0
Impressions (Case3)	3	3	3	3	3	3	2

The EEP adoption rate profiles over the 28 days are shown in Fig. 5.8 and the daily figures for each profile are listed in Table 5.6.

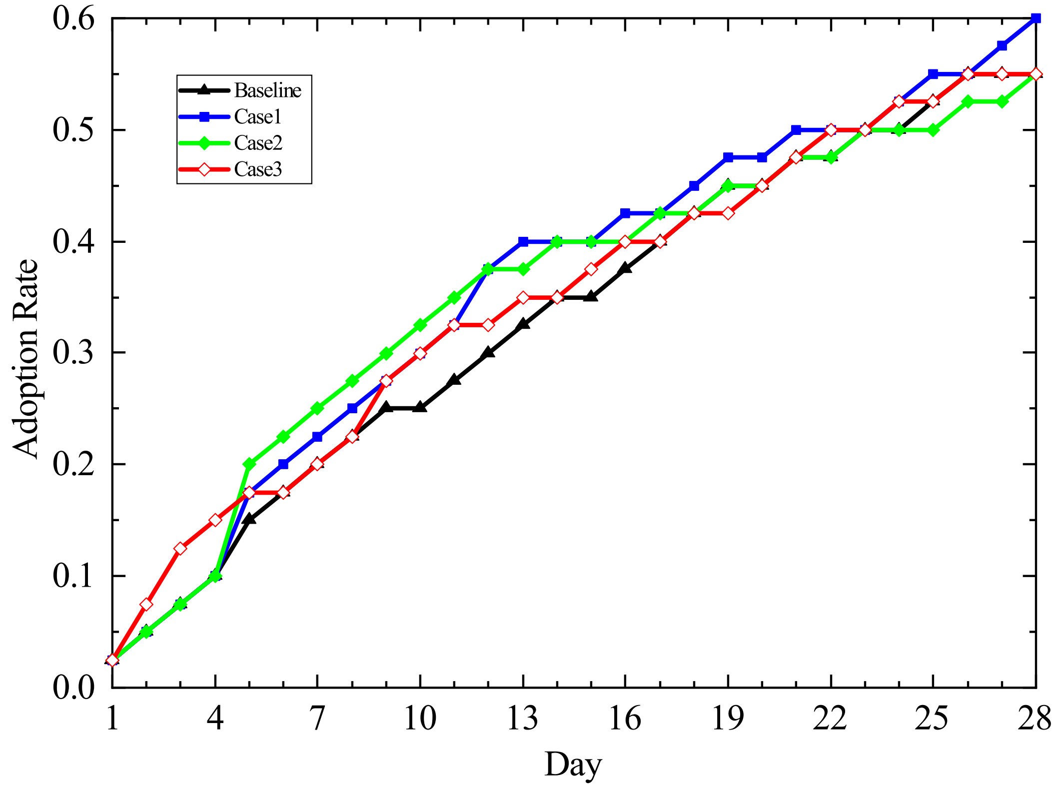


Figure 5.8: Daily Adoption Rate of of Small Network with Different Advertisement Inputs (Baseline, Case 1, Case 2, Case 3)

Inspection of these results reveals that changes in impression volume, $X(t)$, affects the advertisement influence term $ad(t)$, as described in (5.10), which leads to different profiles of the daily adoption rates. Comparing to the baseline, Case 1 exhibits an average increase of 9.09% in adoption rate on Day 28, which is expected as the total amount of advertisement input is larger than the baseline input. In Case 2, the daily adoption rate experiences a rapid growth during the initial 14 days likely driven by the big input in the first week, then the growth slows down, eventually falling slightly below the baseline after Day 26. Case 3 takes a nearly uniform distributed input, the adoption rate profile doesn't seem to be affected much by this distributed input, in fact it demonstrates a rapid increase in daily adoption rate during the first three days, but ultimately the result on Day 28 aligns with the baseline result on the same day.

Table 5.6: Daily Adoption Rate of of Small Network with Different Advertisement Inputs (Baseline, Case 1, Case 2, Case 3)

Day	1	2	3	4	5	6	7
Adoption Rate (Baseline)	0.025	0.050	0.075	0.100	0.150	0.175	0.200
Adoption Rate (Case1)	0.025	0.050	0.075	0.100	0.175	0.200	0.225
Adoption Rate (Case2)	0.025	0.050	0.075	0.100	0.200	0.225	0.250
Adoption Rate (Case3)	0.025	0.075	0.125	0.150	0.175	0.175	0.200
Day	8	9	10	11	12	13	14
Adoption Rate (Baseline)	0.225	0.250	0.250	0.275	0.300	0.325	0.350
Adoption Rate (Case1)	0.250	0.275	0.300	0.325	0.375	0.400	0.400
Adoption Rate (Case2)	0.275	0.300	0.325	0.350	0.375	0.375	0.400
Adoption Rate (Case3)	0.225	0.275	0.300	0.325	0.325	0.350	0.350
Day	15	16	17	18	19	20	21
Adoption Rate (Baseline)	0.350	0.375	0.400	0.425	0.450	0.450	0.475
Adoption Rate (Case1)	0.400	0.425	0.425	0.450	0.475	0.475	0.500
Adoption Rate (Case2)	0.400	0.400	0.425	0.425	0.450	0.450	0.475
Adoption Rate (Case3)	0.375	0.400	0.400	0.425	0.425	0.450	0.475
Day	22	23	24	25	26	27	28
Adoption Rate (Baseline)	0.475	0.500	0.500	0.525	0.550	0.550	0.550
Adoption Rate (Case1)	0.500	0.500	0.525	0.550	0.550	0.575	0.600
Adoption Rate (Case2)	0.475	0.500	0.500	0.500	0.525	0.525	0.550
Adoption Rate (Case3)	0.500	0.500	0.525	0.525	0.550	0.550	0.550

Next the simulation is extended to 180 days for the small network to observe the full system dynamics. Figure 5.9 shows the adoption rate profiles of Baseline and Case 1 for 180 days. These two cases have the same advertisement frequency, once per week, but the magnitude of Case 1 is twice of the Baseline input magnitude. It can be seen that for both input profiles, their steady state values of adoption rate reach the same constant level of 0.9. The double size magnitude for Case 1 doesn't affect the steady state value because both systems approach the saturation status of the adoption rate.

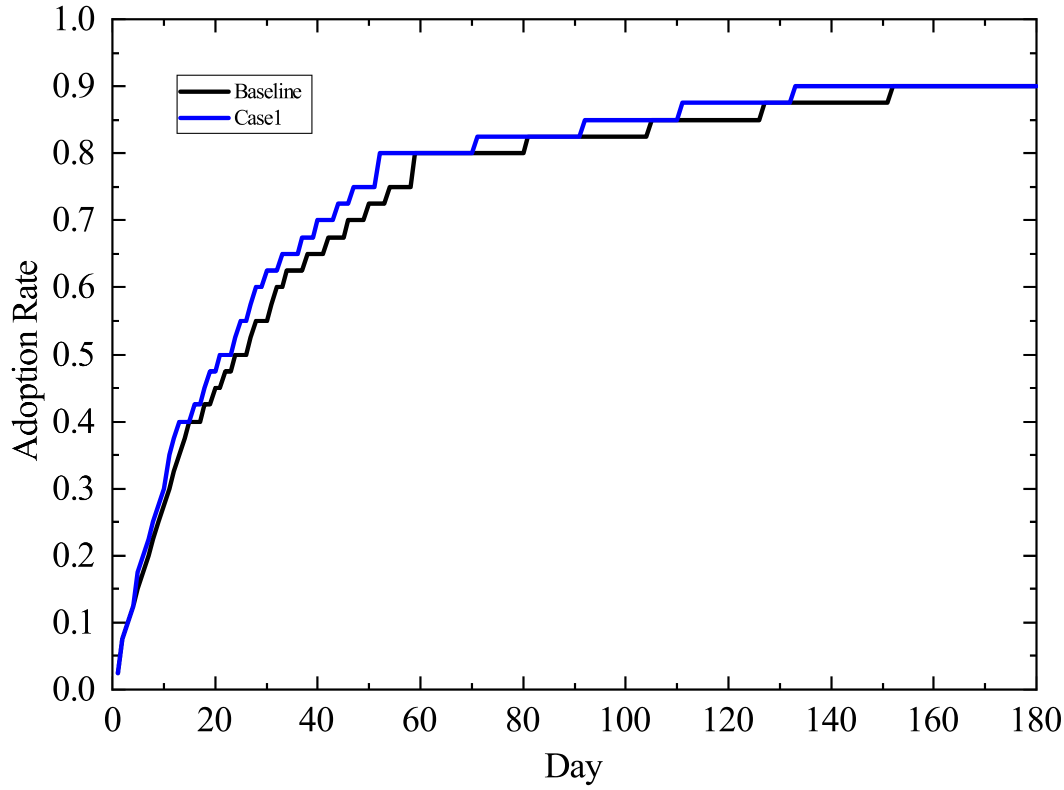


Figure 5.9: Daily Adoption Rate of of Small Network with Different Advertisement Inputs for 180 Days (Baseline, Case 1)

5.3.5 Large Population Network

For the large population social network with one million users/nodes, the network degree is assumed to be 5, that means on average each user is connected to five other users. The complex network model is used to generate the structure of large population network, to be more specific, the W-S small-world network is chosen as the network model since this network model is a classic solution for real world network simulation (Chen et al., 2014).

After the network connections are generated, personal factors are given to each node using the proportions that matches the group distribution according to the survey report (NOP and Research, 2012), figure details are shown in Table 5.7. For a given rewired probability P_r, the pseudo code for large network generation is shown in

Algorithm 5.3.

Table 5.7: Large Network Proportion

Levels (%)	1	2	3	4
Income (a)	30	38	32	-
Children Status (b)	64	36	-	-
Age Group (c)	30	36	13	21
Employment (d)	38	62	-	-

Algorithm 5.3 Generation of Large Network and Node Personal Characteristics

Initialisation of A_{ij} for 1,000,000 nodes with $k = 5$
Initialisation of P_r
while $1 \leq i \leq 1,000,000$ **do**
  for Each source node i **do**
    Generate random personal characteristics a_i, b_i, c_i, d_i according to proportions in Table 5.2, where $a_i \in [1, 2, 3]$, $b_i \in [1, 2]$, $c_i \in [1, 2, 3, 4]$, $d_i \in [1, 2]$
    for Each adjacent node j **do**
      if $A_{ij} == 1$ **then**
        Generate random number $0 \leq rand_{ij} \leq 1$
        if $rand_{ij} \geq P_r$ **then**
          $A_{ij} == 0$
          $A_{ij^*} == 1$, where $j^* \notin [i, j]$
        end if
      end if
    end for
  end for
end while
Output updated A_{ij} and a_i, b_i, c_i, d_i

Once the network is generated, nodes connection status and personal factors for each node are saved for further calculation. The unknown parameters in $\boldsymbol{\theta}_U$ can be estimated with the LS parameter identification algorithm in (5.14). The convergence graph of the nine parameters of the large network is shown in Fig. 5.10 (rewire probability $P_r = 0.45$ for example). It is shown that the parameters entering steady states after 38 iterations and the achieved values are shown in Table 5.8. Among the four weighting parameters in the utility function (5.11), $(\omega_1, \omega_2, \omega_3, \omega_4)$, ω_3 shows the largest value, which suggests the significant influence of the whole network trend on each individual users for EEP adoption decision making.

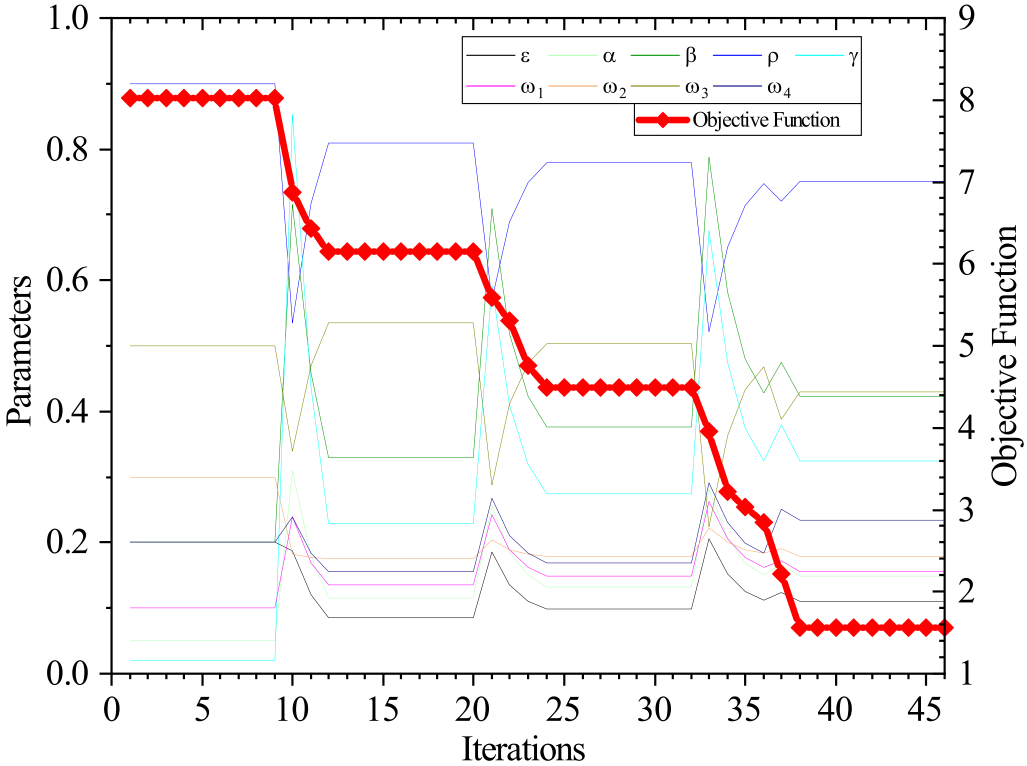


Figure 5.10: Parameter Estimation Profiles for the Large Population Network

Table 5.8: Large Network Model Parameters at $P_r = 0.45$ rewire probability

ε	α	β	ρ	γ
0.1106	0.1487	0.4233	0.7515	0.3238
ω_1	ω_2	ω_3	ω_4	
0.1558	0.1788	0.4304	0.2349	

The calculated adoption rates of the large network and the adoption rates from the survey data are shown in Fig. 5.11. It can be seen that the modelled output stays close to the survey data, which indicates the proposed adoption rate model has a good modelling quality for the large social network under study. Figure 5.11 should ideally show convergence, but since it represents the adoption rate of a large population network, the time period required for convergence should be much longer than 28 days. However, there isn't sufficient advertisement campaign data to support this prolonged observation.

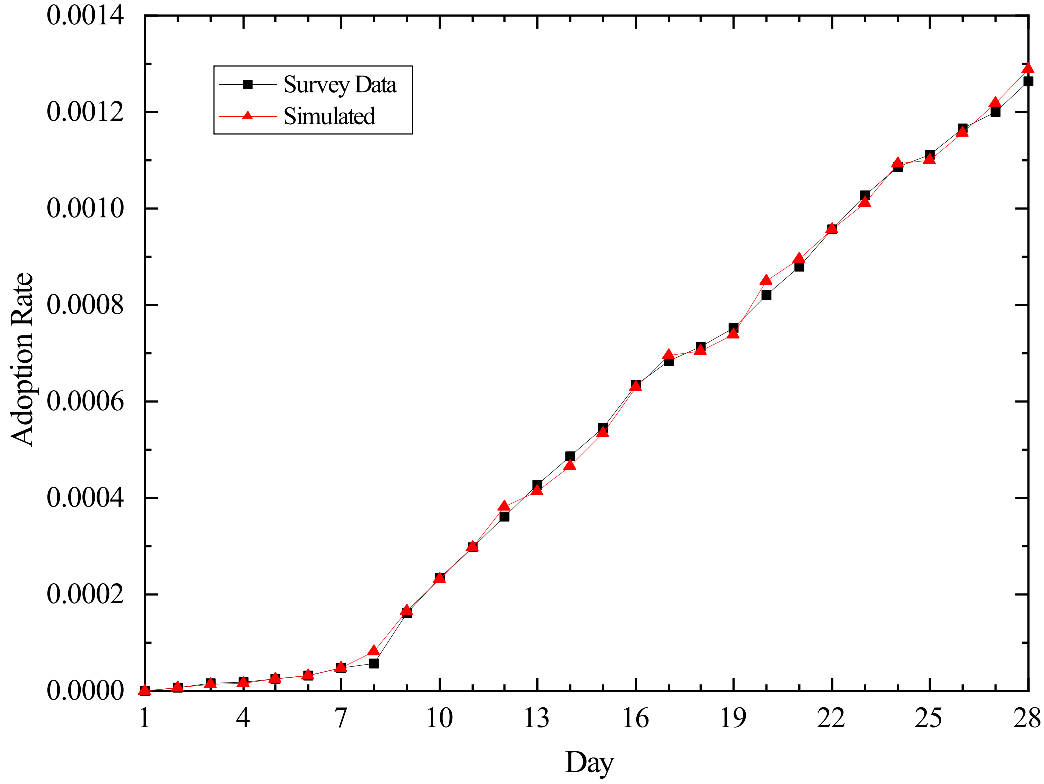


Figure 5.11: Large Network Adoption Rate ($P_r = 0.45$)

For a complex network, its randomness property is determined by the rewire probability parameter, P_r. For the W-S small-world network, a series of P_r values are taken to generate a set of networks representing different levels of randomness from regular network ($P_r = 0$) to E-R random network ($P_r = 1$), with a 0.05 step increase for P_r generation. The results of the four weighting coefficients in the utility function, under the set of P_r, are shown in Fig. 5.12. The data are listed in Table 5.9.

Table 5.9: Large Network Coefficients

Rewire Probability	Coefficients of Utilities			
β_r	ω_1	ω_2	ω_3	ω_4
0	0.0974	0.1447	0.5333	0.2245
0.05	0.1023	0.1464	0.5205	0.2307
0.10	0.1140	0.1766	0.4830	0.2264
0.15	0.1161	0.1598	0.5052	0.2188
0.20	0.1224	0.1576	0.4863	0.2337
0.25	0.1345	0.1505	0.4715	0.2434
0.30	0.1339	0.1643	0.4792	0.2227
0.35	0.1413	0.1515	0.4571	0.2500
0.40	0.1468	0.1513	0.4546	0.2473
0.45	0.1558	0.1788	0.4305	0.2349
0.50	0.1613	0.1691	0.4276	0.2420
0.55	0.1768	0.1699	0.4126	0.2407
0.60	0.1860	0.2087	0.3914	0.2139
0.65	0.1951	0.2048	0.3704	0.2297
0.70	0.2024	0.2101	0.3567	0.2308
0.75	0.2096	0.2100	0.3404	0.2400
0.80	0.2139	0.2266	0.3314	0.2281
0.85	0.2193	0.2165	0.3186	0.2456
0.90	0.2251	0.1970	0.3051	0.2728
0.95	0.2294	0.2075	0.2959	0.2671
1.00	0.2386	0.2051	0.3393	0.2170

It can be seen from Table 5.9 that the social network trend influence has the largest weight (ω_3), which indicates that the network influence utility is more reliant on the EEP adoption rate of the whole social network compared to other factors. Figure 5.12 shows that ω_3 decreases with the increase of the rewire probability. At $P_r = 0$ when the network is clustered, ω_3 takes more than 50% of the total weighting. At $P_r = 1$ when the network becomes an E-R random network, ω_3 drops to about one-third of the total weighting.

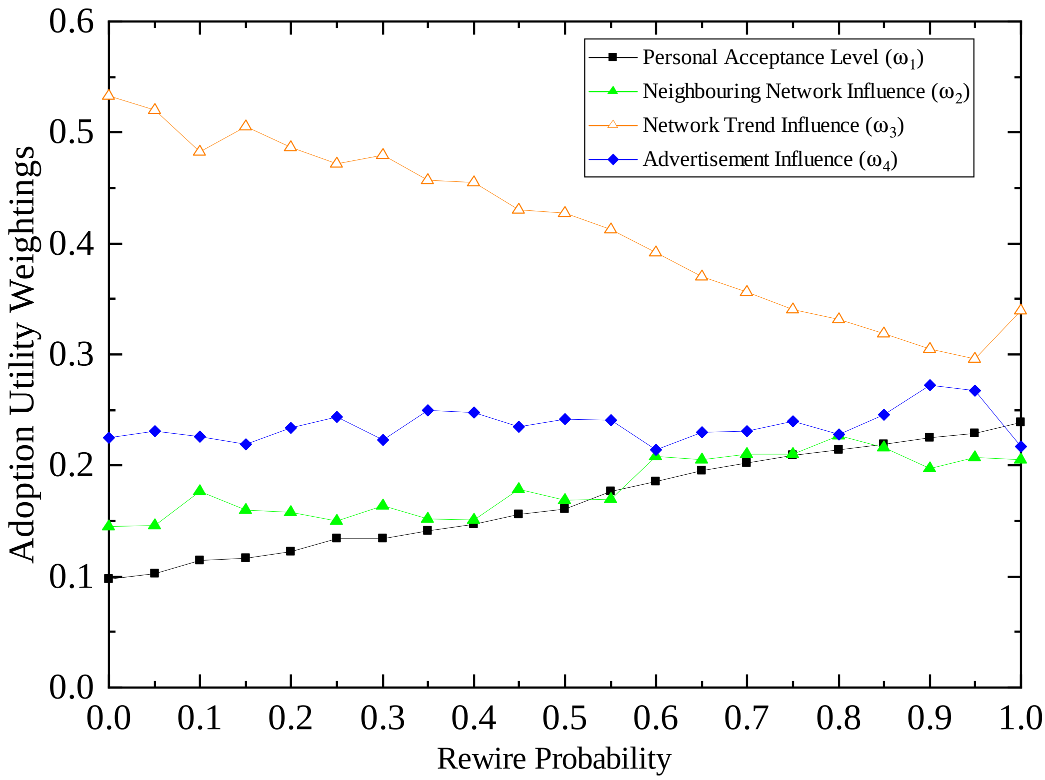


Figure 5.12: Network Influence Utility Weights along with Rewire Probability

The weights of personal acceptance level (ω_1) and neighbour nodes influence (ω_2) increases along with the increase of the rewire probability. That means they have more influence for a large network that is close to a random network than a clustered network. This might due to the reason that a lower rewire probability will generate a network that has different clustered groups in which nodes are more sensitive to group adoption rate. However, for a network that is close to a E-R random network, no such group exists and all nodes are relatively equal compared to clustered network. Thus, the personal acceptance level has more impact in a network with larger rewire probability.

For the advertisement influence, the weight ω_4 varies between 0.2140 to 0.2728, doesn't show a clear pattern of increase or decrease when the rewire probability changes.

Figure 5.13 shows the calculated threshold ε in (5.12) along with the rewire probability. It can be seen that the value of threshold is rather stable with the change of the network structure. Although ε shows a trend of decreasing when P_r is changed from 0 to 1, the value only varies within a small range between 0.11060 to 0.11065.

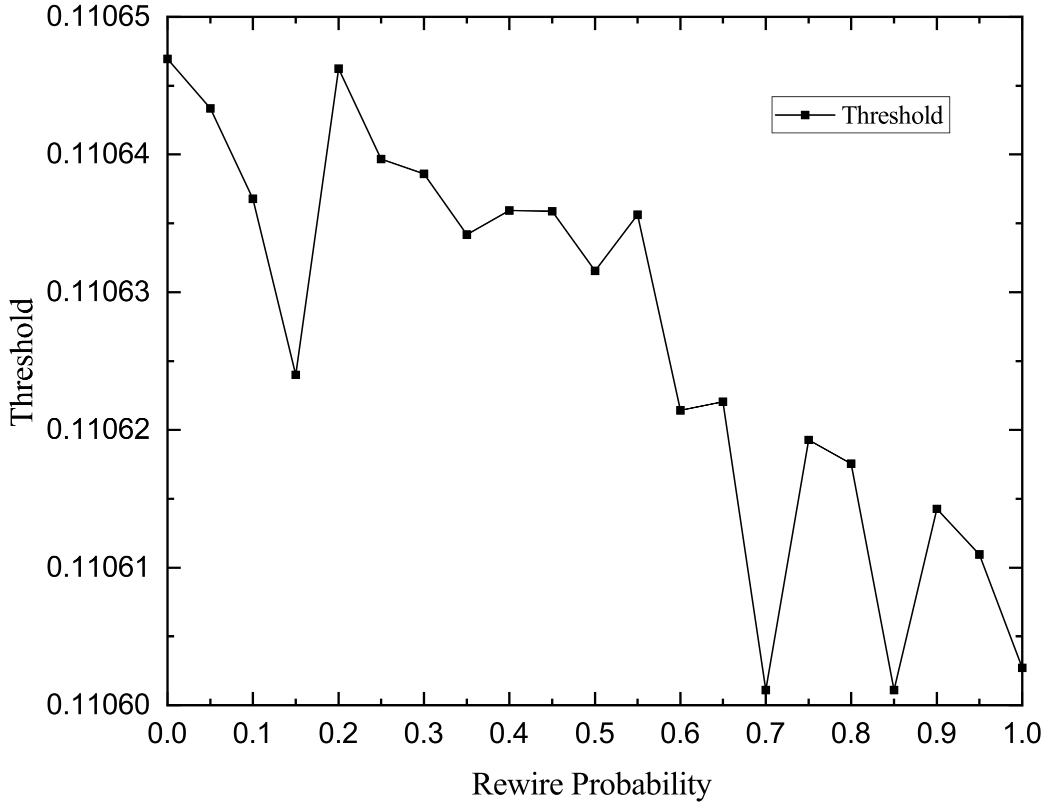


Figure 5.13: Large Network Threshold with Respect to Rewire Probability

5.3.6 Parameter Sensitivity Analysis

In the sensitivity analysis, the parameter values in Table 5.4 are taken as the nominal parameter values for the small network, the parameter values in Table 5.8 are used as the nominal values for the large network model at $P_r = 0.45$. Sensitivity analysis and Monte Carlo calculation are performed for both small and large network.

Similar to the sensitivity analysis in Section 4.4.4, the nine model parameters in $\boldsymbol{\theta}_U$ (5.13) are assumed to follow Gaussian distribution, their mean values are the nominal values, the standard deviations are taken at four levels, i.e., $[0.1, 0.2, 0.5, 1]$ of the mean values. The uncertainty range for each parameter is between zero and twice of the nominal value. In each calculation, one parameter in $\boldsymbol{\theta}_U$ is varied within the uncertainty range and the other parameters are kept at their nominal values.

For each parameter, nine points are uniformly sampled between the uncertainty range. Taking the l_{sa}-th parameter in $\boldsymbol{\theta}_U$ as an example, the 9 sampled points for the parameters are written as

$$\boldsymbol{\theta}_U^{l_{sa}} = [\boldsymbol{\theta}_U^{l_{sa}}(1), \cdots, \boldsymbol{\theta}_U^{l_{sa}}(9)]^\top, \ \ (l_{sa} = 1, \cdots, 9). \tag{5.16}$$

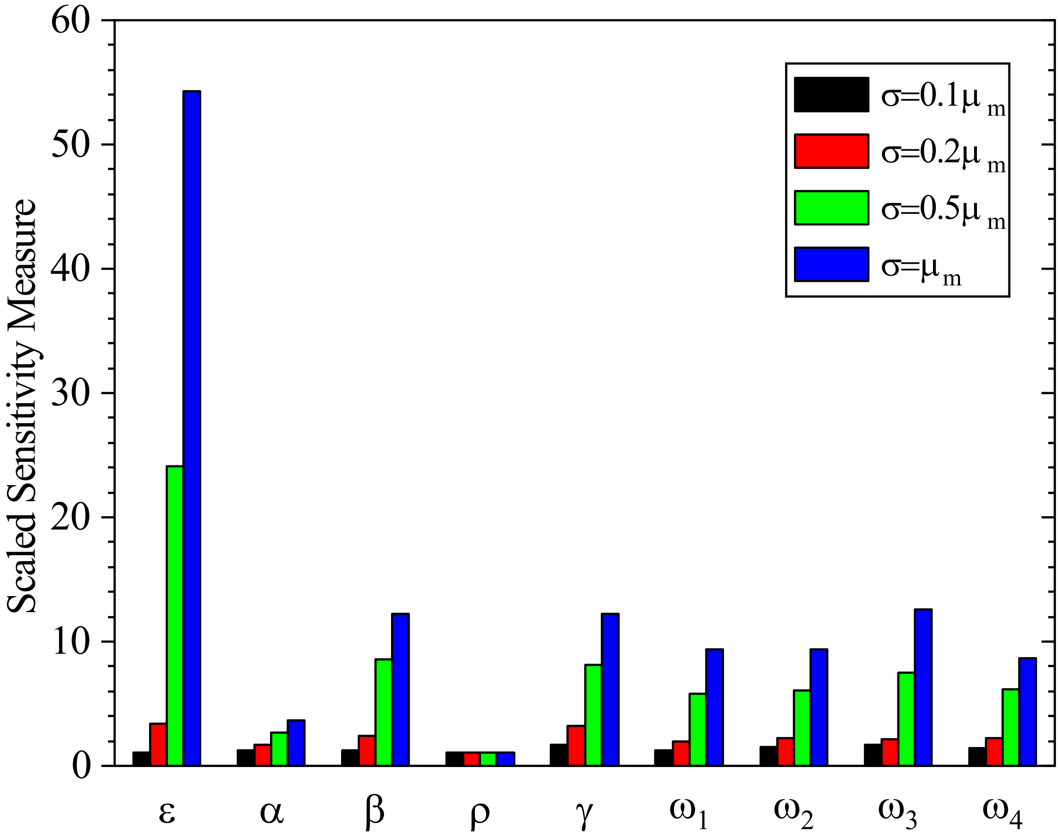


Figure 5.14: Parameter Sensitivity Analysis of Adoption Utility Model (Small Network)

The pdf values corresponding to the nine sampled parameters are written in a vector $\boldsymbol{pdf}_{l_{sa}}$, i.e., $\boldsymbol{pdf}_{l_{sa}} = [\boldsymbol{pdf}_{l_{sa}}(1), \cdots, \boldsymbol{pdf}_{l_{sa}}(9)]^\top$. Then for the l-th sampling point in the l_{sa}-th parameter, the t-th day adoption rate can be calculated as $h(t, \boldsymbol{\theta}_U^{l_{sa}}(l)) \cdot \boldsymbol{pdf}_{l_{sa}}(l)$.

To assess the relative impacts from each of the parameters in $\boldsymbol{\theta}_U$, the following function measuring the diffrence between the calculated adoption rates and the surveyed adoption rates, weighted by sampled parameter pdfs, is used as the sensitivity output for the l_{sa}-th parameter.

$$J_{l_{sa}} = \sum_{t=1}^{28} \sum_{l=1}^{9} \Big(R\big(t, \boldsymbol{\theta}_U^{l_{sa}}(l)\big) - y_R(t) \Big)^2 \cdot \mathbf{pdf}_{l_{sa}}(l) \tag{5.17}$$

For each parameter, taking the standard deviation at four levels, i.e., 10%, 20%, 50% and 100% of the nominal value, and apply the calculation to the nine parameters in θ_U, the sensitivity analysis results are shown in Fig. 5.14 for the small network and Fig. 5.15 for the large network, where the output metric $F_{l_{sa}}$ is scaled by the value calculated at the nominal parameter values.

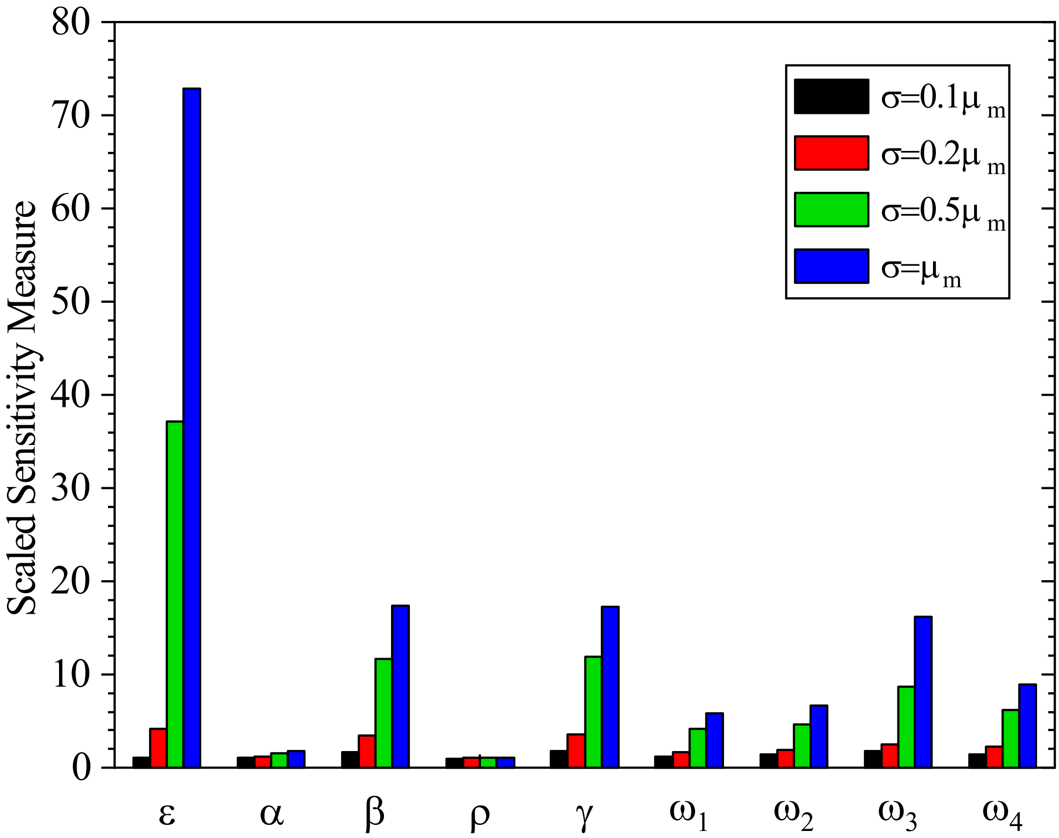


Figure 5.15: Parameter Sensitivity Analysis of Adoption Utility Model (Large Network)

According to the parameter sensitivity analysis results, among the nine parameters in $\boldsymbol{\theta}_U$, the threshold for node state change ε has the strongest influence on the formed output function in (5.17). In the group of parameters for the advertisement model, $(\alpha, \beta, \rho, \gamma)$, the diminishing return factor γ and the upper threshold β have larger sensitivity. For the four weighting coefficients $(\omega_1, \omega_2, \omega_3, \omega_4)$, their influence are at similar levels for the small network, but for the large network, the weighting coefficient on network trend (ω_3) has the largest influence to model output.

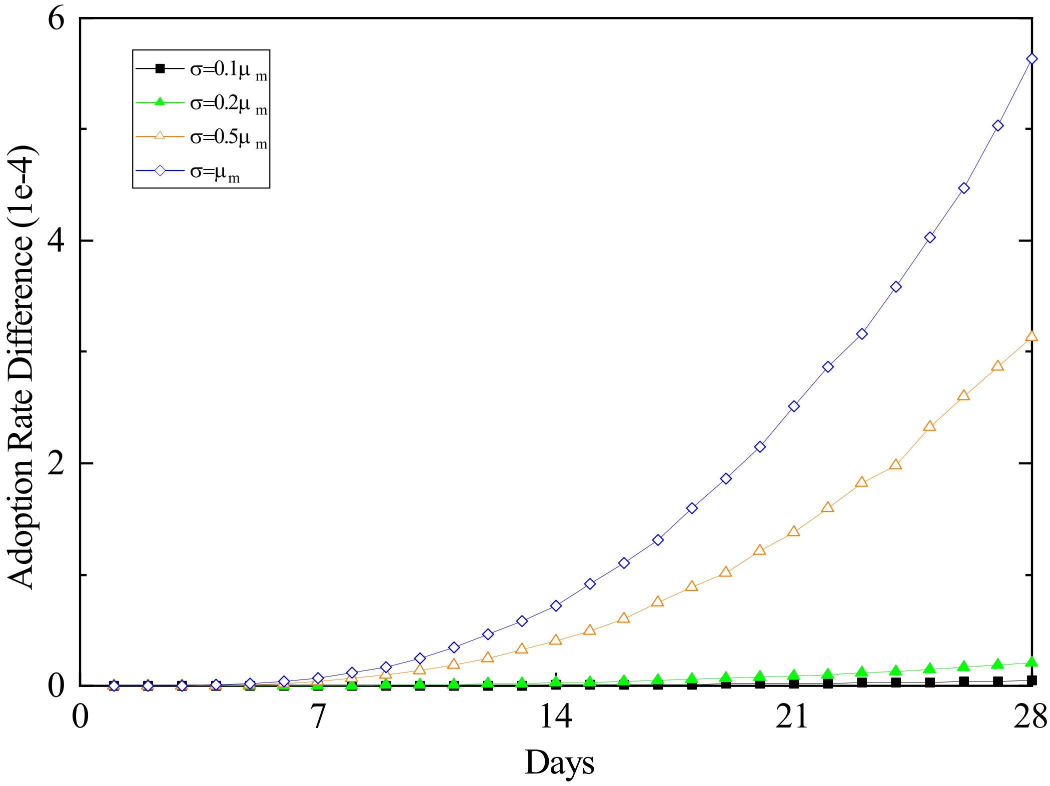


Figure 5.16: Difference of Adoption Rates in Large Population Network Achieved by Nominal Value and Monte-Carlo Calculation under Parameter Variation (Parameter ε, Standard Deviations: 10%, 20%, 50% and 100% of Mean Value)

Next, the node change threshold ε is chosen for further calculations under four different standard deviations. Again, the same sampling of nine points is applied to ε, and the calculated adoption rates are weighted by pdf values. The daily adoption rates calculated by this Monte-Carlo method is compared against the adoption rates achieved by nominal value, the difference between these two are shown in Fig. 5.16, from which it can be seen that when the parameter uncertainty is larger (with larger standard deviation), the difference between the model result and the surveyed data is larger as expected. The difference caused by model uncertainty is quite large, reaching 44% at 100% standard deviation, for over the 28 days. This indicates that the model is sensitive to the value of adoption threshold (ε).

5.4 Summary

In this chapter, a mathematical model has been developed to describe the dynamics of the EEP adoption rate for both small and large population networks. For the small network, the network connections are modelled using the survey data. For the large network, the complex network model, W-S small world model, is employed to generate the network, for which the randomness level depends on the rewire probability.

A network influence utility function is defined to include the four factors that are considered to impact the EEP adoption rate: the personal acceptance level, the neighbour nodes social influence, the overall network trend influence, and the advertisement influence. The utility coefficients associated with the four factors are calculated based on survey data and the proposed adoption rate models using LS estimation.

Both small and large population networks are simulated in case studies. Comprehensive analysis has been made to examine the proposed model from different perspectives. Sensitivity analysis are conducted to understand the influence of parameter uncertainties and the relative influence between multiple factors. For the large network, the rewire probability is varied from 0 to 1 to see how the randomness of network structure affects the model output. Several advertisement input patterns are applied to the small network to compare their influences to the EEP adoption rate. These analyses show that the proposed model is robust to model uncertainty and can be used to analyze influence from multiple factors under different network structures.

The personal acceptance depends on various factors, among them the family status and the employment status have most influence to the decision making on EEP adoption. For both small and large population networks, the overall network trend has the most impact in the network utility function, which indicates that the most influential factor to the EEP adoption behaviour for a householder is the adoption rate of the whole social network. For the small population network under study, the personal acceptance level accounts for 23.88% among four weights in the utility function. For the large population network under study, the personal acceptance level shows less impact when the network is close to a clustered network. However, with the increase of rewire probability, when the network is approaching an E-R random network, the influence of the personal acceptance level counts more, 23.86% in the whole weighting. According to (Chen et al., 2014), real world social network is close to a W-S small-world model that has rewire probability around 0.1. Thus, personal acceptance has less impact for a large population network compared to a small population network.

From the model-based analysis in this chapter, the influence of advertisement input (measured by impression) towards the EEP adoption rate is examined as a preliminary study. Further investigation on influence of advertisement control to energy savings will be preserensted in Chapter 6.

Chapter 6

Influence of Advertisement Control to Energy Savings

6.1 Introduction

In a social network, individual users can be considered as nodes, and those nodes that carry energy savings information are defined as hosts or information source nodes. Similar to the epidemics theory, when hosts interact with other nodes in the network, information starts to diffuse. In a mass roll-out programme of EEP, the initial households with products installed can be seen as hosts. These hosts that benefit from the EEP may spread the information among their neighbouring nodes. The latter that later install the product then become sub-hosts. It is noted that even if some neighbour nodes do not install the EEP, they still carry energy saving product information that can be further spread to wider neighbours. In this way, more adoption of energy saving products are expected to be achieved through the interactions between individuals within the network.

For a mass roll-out programme, there is always a product installation or cost savings target to be achieved within a given time period. During the implementation process of the programme, the actual product roll-out may be slower than expected. In this situation, household engagement measures such as advertisement can be taken to increase the adoption of EEP. This process of increasing energy savings depends not only on user behaviour but also on the EEP settings. The EEP needs to be profitable for an effective long-term goal. Traditionally, energy supply companies meant to increase product profit either by expanding market scale or by decreasing unit product cost (George et al., 1970). The EEP cost setting is constrained by technical limitations and requirements on product quality. Therefore, expanding market is regarded as a favoured option in many energy saving schemes.

It is proved that increased advertisement has a positive association to the improved market share (Shapiro, 1983). The rapid development of modern media technologies leads to various means of advertising, such as TV advertisement, newsletters, emails, social media, internet website and online video streaming, etc. Hence, it is crucial to understand the effectiveness of advertisements to promoting EEP, and particularly those popular online advertisements through internet. In online advertisement business, standardised units of advertisement are called impressions. An impression refers to the exact times that the advertisement has been viewed. The increasing number of purchasing EEP is expected to be affected by impression. A crucial and challenging task is to quantify the relationship between the impression and the product adoption rate, which hasn't been reported in any literature as far as we know. This motivates the development in this chapter.

The influence of user behaviour to energy savings in social networks are studied in Ekpenyong et al. (2014), where concepts of direct and indirect energy savings are introduced and used in calculation of the total EES of a small network. The model developed in our earlier work (Du et al., 2016) and in Chapter 4 can be used to simulate the diffusion of energy savings information within a small population network (40 individuals in the case study), the nodes connections and strength are obtained by survey data. These studies rely on detailed network connection information, suitable for small scale networks only. To the best of our knowledge, none of the existing models considers the advertisement factor which has a clear influence on user behaviour and will eventually affect the promotion of new energy saving technologies.

In this work, the influence of online advertisement to energy saving programme is studied through modelling of EES and advertisement control based on the BFT (West and Harrison, 1997). The proposed model includes the information diffusion model, the advertisement model and the electricity cost savings calculation, where the key input is the daily advertisement rate and the output is the increased adoption rate due to the influence of advertisement. In order to describe the information diffusion dynamics for a large-scale network, two methods from complex network theory, the epidemics theory (Chen et al., 2014) and the W-S small-world network theory (Watts and Strogatz, 1998), are employed to establish the model. With this model established, the energy savings due to the advertisement influence in the social network can be quantitatively determined.

The remainder of the chapter is structured as follows. A novel model of EES that includes information diffusion and advertisement input is established in Section 6.2 to quantify the energy savings influenced by advertisement control in a large social network. An advertisement control optimisation method is proposed in Section 6.3 considering several design objectives. The case studies are conducted for small and large networks, with results and discussions presented in Section 6.4. Conclusions and a summary are given in Section 6.5. The campaign data for the large population network case studies are given in the appendix.

6.2 Advertisement Control Model for Adoption Rate and Energy Savings

In the previous chapter, a model for EES calculation is established for a small network using probability theory (see Section 4.2.5), where the advertisement input is not included. The EES model in Chapter 5 mainly applies to small networks when full connection information between nodes is available. For a large network considering advertisement input, alternative methods are required for modelling. In this section, a new mathematical model is developed to calculate the EES dependence on advertisement influence and network interactions. Different from the model in Du et al. (2016) and in Section 4.2.5 that uses network connection details between individuals, the information diffusion model in this work is established based on the epidemics theory, which does not rely on individual connection information, thus can be applied to large population networks. The advertisement input is added to the epidemics theory framework.

6.2.1 Epidemic Model

Central to the study of epidemic models is the SIR model (Chen et al., 2014), which categorizes a population into three distinct compartments: susceptible individuals (S), infected individuals (I), and those who have recovered (R). The balance between these compartments at any given time t is illustrated by (6.1):

$$S(t) + I(t) + R(t) = 1. \tag{6.1}$$

This model tracks the shift of individuals among the three compartments, enabling tools to analyze the dynamics of virus spread within a population and assess various intervention measures' effectiveness. Here, the susceptible group comprises individuals vulnerable to infection but currently uninfected. The infected segment encompasses those currently battling the disease and potentially transmitting it. The recovered segment comprises individuals who have either overcome the disease or succumbed to it, no longer contributing to its transmission.

Grounded on a series of equations, the SIR model delineates the transition flow between these compartments, including variables such as the infection transmission rate from infected to susceptible individuals, symbolized by v, and the recovery or death rate of infected individuals, represented by δ. By resolving these equations, the disease progression within a population over time and its ultimate state can be simulated. The variation of each state can be articulated in (6.2) (Chen et al., 2014):

$$\begin{aligned} \tfrac{\mathrm{d}S(t)}{dt} &= -vS(t)I(t),\ S(0) = S_0 \\ \tfrac{\mathrm{d}I(t)}{dt} &= vS(t)I(t) - \delta I(t),\ I(0) = I_0, \\ \tfrac{\mathrm{d}R(t)}{dt} &= \delta I(t),\ R(0) = 0, \end{aligned} \tag{6.2}$$

where the initial compartment states are denoted as S_0, I_0, and R_0.

6.2.2 Information Diffusion Model without Advertisement Input

Similar to the epidemic model, for information diffusion, those unacquainted with specific information can be deemed susceptible. Those informed and acting upon said information equate to the infected and recovered nodes, respectively. The SIR model's appeal lies in its streamlined nature, facilitating large population network simulations without delving into the details of network structures. This feature renders it a useful tool for modelling, offering insights into information spread and evaluation of assorted intervention strategies.

Consider a network with N nodes, and each node in the network can only have one state at any given time instant. Following the idea of the SIR model, the states of nodes in energy networks can be defined as follows:

- users unaware of the energy efficiency information, represented by S;
- users aware of the energy efficiency information, represented by I; and
- users purchased the product, represented as the EEP adopted state R.

In a size-N network at time t, the three states are represented by $S(t)$, $I(t)$ and $R(t)$, respectively. The sum of $S(t)$, $I(t)$ and $R(t)$ is equal to one at all times, which is the same as in (6.1).

Define the probability to transit from the unaware state, S, to the aware state, I, to be ν, and the probability from the aware state, I, to the adopted state, R, to be δ. It is assumed that ν and δ are constant parameters taking values between 0 and 1. In the SIR model (Chen et al., 2014), a node in state R will not interact with other nodes in the network. Thus, the changing rate of $S(t)$ only depends on state $S(t)$ and state $I(t)$. However, in the energy social network model, individual users that have already adopted the EEP may still interact with unaware users and pass on energy savings information. Therefore, the changing of $S(t)$ in the energy network should consider interactions from nodes under both state I and state R. Thus, the dynamics of the three states is represented by the following set of ordinary differential equations:

$$\begin{aligned}\frac{\mathrm{d}S(t)}{\mathrm{d}t} &= -\nu S(t)\big(I(t)+R(t)\big)\\ \frac{\mathrm{d}I(t)}{\mathrm{d}t} &= \nu S(t)\big(I(t)+R(t)\big)-\delta I(t)\\ \frac{\mathrm{d}R(t)}{\mathrm{d}t} &= \delta I(t)\end{aligned} \tag{6.3}$$

Following the small-world network theory, denote k as the average network degree of the considered network, then the probability parameters ν and δ in (6.3) can be represented as

$$\nu = k\psi, \quad \delta = k\varphi, \tag{6.4}$$

where ψ is the probability of state transition from S to I, and φ is the probability of state transition from I to R. The three parameters, k, ψ and φ are considered as constant. For a network with a large population, the initial value of the adopted state R can be assumed to be zero because the information sources only occupy a negligible small section of the whole population. Therefore, $S(0) = S_0$, $I(0) = I_0$ and $R(0) = 0$, where S_0 and I_0 are the initial states for S and I. With the use of the average network degree term, the model in (6.3) can be written as

$$\begin{aligned}\frac{\mathrm{d}S(t)}{dt} &= -k\psi S(t)\big(I(t)+R(t)\big)\\ \frac{\mathrm{d}I(t)}{\mathrm{d}t} &= k\psi S(t)\big(I(t)+R(t)\big)-k\varphi I(t)\\ \frac{\mathrm{d}R(t)}{\mathrm{d}t} &= k\varphi I(t).\end{aligned} \tag{6.5}$$

Equation (6.5) is the information diffusion model for energy saving networks without including the advertisement impact. The following discrete-time model is used in simulation implementation.

$$\begin{aligned}S(\tau) &= S(\tau-1) - k\psi S(\tau)\big(I(\tau)+R(\tau)\big)\cdot\Delta\tau,\\ I(\tau) &= I(\tau-1) + \Big(k\psi S(\tau)\big(I(\tau)+R(\tau)\big)-k\varphi I(\tau)\Big)\cdot\Delta\tau,\\ R(\tau) &= R(\tau-1) + k\varphi I(\tau)\cdot\Delta\tau.\end{aligned} \tag{6.6}$$

where τ is discrete time index, $\Delta\tau$ is the sampling period which is taken as one day in this study. The discrete-time form is obtained by Euler's Backward differencing method with sampling time being 1.

6.2.3 Information Diffusion Model with Advertisement Input

When consider advertisement influence alone, the dynamics of $S(t)$, $I(t)$, and $R(t)$ can be described as

$$\begin{aligned} \tfrac{\mathrm{d}S(t)}{\mathrm{d}t} &= -\omega(t)S(t) - \mu(t)S(t) \\ \tfrac{\mathrm{d}I(t)}{\mathrm{d}t} &= \omega(t)S(t) - \mu(t)I(t) \\ \tfrac{\mathrm{d}R(t)}{\mathrm{d}t} &= \mu(t)\big(S(t) + I(t)\big) \end{aligned} \tag{6.7}$$

where $\mu(t)$ is the advertisement adoption rate, $\omega(t)$ is the advertisement positive response rate. Integrating the advertisement model in (6.7) with the original information diffusion model in (6.5), the network adoption dynamics is written as the following differential equation model:

$$\begin{aligned} \tfrac{\mathrm{d}S(t)}{\mathrm{d}t} &= -k\psi S(t)\big(I(t) + R(t)\big) - \omega(t)S(t) - \mu(t)S(t), \\ \tfrac{\mathrm{d}I(t)}{\mathrm{d}t} &= k\psi S(t)\big(I(t) + R(t)\big) - k\varphi I(t) + \omega(t)S(t) - \mu(t)I(t), \\ \tfrac{\mathrm{d}R(t)}{\mathrm{d}t} &= k\varphi I(t) + \mu(t)\big(S(t) + I(t)\big) \end{aligned} \tag{6.8}$$

with $S(0) = S_0$, $I(0) = I_0$ and $R(0) = R_0$. In this way, the advertisement control is integrated in the network model, and its time-dependent influence to the adoption rate can be described. Since the advertisement operation is run in discrete time, the continuous-time model in (6.8) is discretized into the following model using the Euler's backward differencing method with sampling time being 1.

$$\begin{aligned} &S(\tau) = S(\tau-1) - k\psi S(\tau)\big(I(\tau) + R(\tau)\big) + S(\tau)\big(\omega(\tau) + \mu(\tau)\big), \\ &I(\tau) = I(\tau-1) + k\psi S(\tau)\big(I(\tau) + R(\tau)\big) - k\varphi I(\tau) + \omega(\tau)S(\tau) - \mu(\tau)I(\tau), \\ &R(\tau) = R(\tau-1) + k\varphi I(\tau) + \mu(\tau)\big(S(\tau) + I(\tau)\big). \end{aligned} \tag{6.9}$$

Note that the two parameters in the advertisement model, $\omega(\tau)$ and $\mu(\tau)$, are time varying and need to be updated at each time instance during the dynamic system simulation. The calculation of these two time-dependent parameters is discussed in the next subsection.

6.2.4 Update of Time-dependent Parameters in Advertisement Model

At time τ, the parameter $\mu(\tau)$ is the rate of product adoption affected by the advertisement impression (number of advertisement views), denoted by $X(\tau)$. Similarly, the advertisement positive response rate (positive to advertisement), $\omega(\tau)$, is the ratio of the number of users who have positive response to the advertisement over the impression. To be more specific, following an online advertisement, the purchasing behaviour is taken as an action of adoption, the action of clicking the advertisement is taken as a positive response. A positive response does not necessarily lead to adoption directly, however, it provides a chance of information diffusion through the network. In this work, the number of users giving positive response does not include those who have already adopted the EEP before viewing the advertisement.

In the following, the BFT in West and Harrison (1997) is employed to calculate the advertisement adoption rate $\mu(\tau)$ and the positive response rate $\omega(\tau)$. As in (5.9), define $E(\tau)$ to be the effect of impression $X(\tau)$ imposed on the adoption rate, the effect of impression at time τ can be calculated, i.e.

$$E(\tau) = (\beta - \alpha) - \big(\beta - \alpha - \rho E(\tau - 1)\big) \cdot \exp\big(-\gamma X(\tau)\big) \tag{6.10}$$

The advertisement adoption rate is then written as

$$\mu(\tau) = \alpha + E(\tau) = \beta - \big(\beta - \alpha - \rho E(\tau - 1)\big) \cdot \exp\big(-\gamma X(\tau)\big) \tag{6.11}$$

Similarly, the advertisement positive response rate, $\omega(\tau)$, can be determined by

$$\omega(\tau) = \beta_\omega - \big(\beta_\omega - \alpha_\omega - \rho_\omega E_\omega(\tau - 1)\big) \cdot \exp\big(-\gamma_\omega X(\tau)\big) \tag{6.12}$$

The subscript 'ω' is introduced in (6.12) to distinguish parameters for the positive response rate from those for the adoption rate.

6.2.5 Estimation of Model Parameters

The model parameters for energy product information diffusion and advertisement impact are unknown. In this work, we use survey data and LS system identification method to estimate the parameter values. Denote

$$\boldsymbol{\theta} = \left[\begin{array}{cccccccccc} \psi & \varphi & \alpha & \beta & \rho & \gamma & \alpha_\omega & \beta_\omega & \rho_\omega & \gamma_\omega \end{array} \right]^\top \tag{6.13}$$

The initial values are set up as follows. In the information diffusion model, $\psi_0 = 0.1$ or 0.01, $\varphi_0 = 1e - 4$ or $1e - 5$, for smaller or larger population network respectively, $\alpha_0 = \alpha_{\omega 0} = 0$, $\beta_0 = \beta_{\omega 0} = 0.85$, $\rho_0 = \rho_{\omega 0} = 0.9$, $\gamma_0 = \gamma_{\omega 0} = 0.02$. The values of $\boldsymbol{\theta}$ can be calculated by

$$\boldsymbol{\theta^*} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} J_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \tag{6.14}$$

where

$$J_{\boldsymbol{\theta}} = \sum\nolimits_{\tau=1}^{M} \big(R(\tau,\boldsymbol{\theta}) - y_R(\tau)\big)^2 \tag{6.15}$$

is the residual function to be minimised for parameter estimation, M is the length of data used for modelling, Θ is the searching domain for $\boldsymbol{\theta}$, which is a set of positive real numbers in the range of [0,1], $y_R(\tau)$ is adoption rate data obtained from survey, $R(\tau, \theta)$ is the adoption rate data calculated by the model. The parameter estimation algorithm is implemented in MATLAB. The GA algorithm is used to find the optimal solution.

6.2.6 Electricity Cost Saving Model

In this subsection, a cost saving calculation model is derived for decision makers managing a mass roll-out programme. The time index τ is taken for a sampling rate of one day. The EES is used to quantify the energy savings, which can be calculated for the whole network from time 1 to time T as follows:

$$EES_T = N \cdot \sum\nolimits_{\tau=1}^{T} R(\tau) \cdot E, \tag{6.16}$$

where N is the network population size, E is the amount of energy saved by one piece of EEP compared to no product being adopted, and $R(\tau)$ is the number of product adoption at time τ, which is calculated by the information diffusion model in (6.9). Assume that there are z free trial products given to the network, then

$$S(0) = 1 - (z/N), \;\; I(0) = 0, \;\; R(0) = z/N. \tag{6.17}$$

The pseudo-code of EES calculation can be seen in Algorithm 6.1,

Algorithm 6.1 EES Calculation Using Advertisement Control

Initialise network SIR states S_0, I_0, R_0, see equation (6.17)
Initialise advertisement control coefficients, μ and ω
for $1 \leq \tau \leq T$ **do**
    Calculate adoption rate due to advertisement $\mu(\tau)$ see equation (6.11)
    Calculate positive response rate $\omega(\tau)$ see equation (6.12)
    Calculate each states of S, I and R see equation (6.9)
end for
Calculate EES at time T EES_T see equation (6.16)
Output EES_T

Consider the fixed electricity tariff at £π per kWh, the total amount of saved cost from energy savings is

$$C_T = \pi \cdot EES_T = \pi \cdot N \cdot \sum\nolimits_{\tau=1}^{T} R(\tau) \cdot E \tag{6.18}$$

6.2.7 Advertisement Cost and Free Trial Product Cost

For the large population network, assume that a project is set up to achieve the target EEP adoption rate over a given period of time. Define the total budget to be £B, the electricity tariff to be £$\pi(\tau)$ per kWh on Day τ, the price of one million impressions to be $\tilde{\tau}$. The cost of advertisement on day τ is $C_{adv}(\tau) = \tilde{\tau} X(\tau)$. The cost of advertisement from Day 1 to Day T is calculated by

$$C_{adv} = \tilde{\tau} \sum\nolimits_{\tau=1}^{T} X(\tau) \tag{6.19}$$

For the large network, the total cost is equivalent to the advertisement cost.

For small population networks, free trial products are applied. Assume that the cost of one piece of free trial product to the roll-out organiser is C_0, the cost for free trial products can be calculated as zC_0, the total cost for EEP promotion is therefore

$$C_{total} = zC_0 + \tilde{\tau} \sum\nolimits_{\tau=1}^{T} X(\tau). \tag{6.20}$$

6.3 Advertisement Control Design

In real world applications, the objective of an energy saving product roll-out programme should focus on the balance between multiple measures such as the adoption rate, the time period applied, the total cost savings, etc. The following optimization objectives are proposed in this work: (i) minimize the total cost of advertisement, (ii) minimize the time (cost), (iii) maximize the total energy savings, and (iv) maximize the network adoption rate. The advertisement strategy is designed using an optimization algorithm.

The control variable is the advertisement daily impression, $X(\tau)$. Denote

$$\mathbf{X} = \left[X\left(M+1\right), X\left(M+2\right), \ldots, X\left(M+K\right)\right]^{\top} \tag{6.21}$$

as the time sequence vector of impressions over K days, M is the number of days used for modelling, therefore $(M+1)$ is the starting day for the optimization design. The design of the daily impressions can be formulated as a general optimization problem, i.e.,

$$\boldsymbol{X}^{*} = \arg \min_{\boldsymbol{X} \in Z_X} J(\boldsymbol{X}) \tag{6.22}$$

where J is the cost-related objective function to be minimized for each scenario in the case study, the set Z_X is the searching domain for $\boldsymbol{X}$.

- Method 1: targeting the lowest advertisement cost

  The objective of this method is to reach the target adoption rate at the end of the programme with the lowest advertisement cost. Two constraints need to be applied, one is to keep the total cost under the given budget, i.e.,

$$C_T\left(\boldsymbol{X}\right) + P_0 \leq B \tag{6.23}$$

where $C_T(\boldsymbol{X})$ is calculated by (6.19), P_0 is the price of free trial product, B is budget limit. The other constraint is that the adoption rate at the end of the programme must reach the target rate, i.e.,

$$R\left(M+K, \boldsymbol{X}\right) \geq \frac{N_g}{N} \tag{6.24}$$

$R(M+K)$ is the adoption rate at the end of the programme, N_g/N is the target adoption rate to be reached for the network. The optimisation problem is written as

$$\begin{gathered} \boldsymbol{X}^* = \arg \min_{\boldsymbol{X} \in Z_X} C_T(\boldsymbol{X}) \\ \text{subject to (6.23) and (6.24)} \end{gathered} \tag{6.25}$$

- Method 2: targeting the shortest time to achieve required adoption rate

  The objective is to be reach the target adoption rate within the shortest time period under the given budget. This is achieved through the following optimization route

$$\begin{gathered} \boldsymbol{X}^* = \arg \min_{\boldsymbol{X} \in Z_X} T_s(\boldsymbol{X}) \\ \text{subject to (6.23) and (6.24)} \end{gathered} \tag{6.26}$$

  where T_s is the time to achieve the required adoption rate, it is calculated from the numerical procedure by recording the time taken when the required EEP adoption rate is reached.

- Method 3: targeting the largest energy savings

  The objective is to achieve the largest electricity cost savings while reaching the target adoption rate within the budget and time. To achieve the maximum EES_T, calculated by (6.16), the optimization problem is written as follows

$$
\begin{aligned}
\boldsymbol{X}^* = \arg \min_{\boldsymbol{X} \in Z_X} & -EES_T(\boldsymbol{X}) \\
\text{subject to } & (6.23) \text{ and } (6.24)
\end{aligned}
\tag{6.27}
$$

- Method 4: targeting the largest installation number

  This design aims to achieve the largest adoption rate within the given budget and time. The adoption rate is calculated from the information diffusion model in (6.9). The objective function is formulated as follows.

$$
\begin{aligned}
\boldsymbol{X}^* = \arg \min_{\boldsymbol{X} \in Z_X} & -N \cdot R(M+K, \boldsymbol{X}) \\
\text{subject to } & (6.23) \text{ and } (6.24)
\end{aligned}
\tag{6.28}
$$

The above advertisement control design procedure can be summarized into five main steps as the following algorithm.

Algorithm 6.2 Advertisement Control Design Procedure

1: Initialize model parameters in $\boldsymbol{\theta}$ and the initial states of S, I and R.
2: Determine the ten parameters in $\boldsymbol{\theta}$, (6.13) using survey data and the LS identification algorithm in (6.14).
3: Calculate the two time-dependent parameters in the advertisement model (6.9), $\omega(\tau)$ and $\mu(\tau)$, using the BFT in (6.10) - (6.12). Here the model parameters calculated from Step (ii) are used and the calculation applies to a time sequence.
4: Calculate the energy savings as in (6.16), the advertisement cost C_T in (6.19), the adoption rate by model (6.9).
5: Apply the optimisation design for different scenarios in (6.25), (6.26), (6.27) and (6.28) subject to budget and time constraints in (6.23) and (6.24).

6.4 Case Studies

Two case studies are investigated in this work, one is for a small population social network, used in Chapter 4 and Chapter 5, the other one for a large population social network, introduced in Chapter 5.

6.4.1 Experimental Settings and Parameter Estimation

1. Assumptions for the small population network

   - Assume that at the beginning of the programme only one individual is given free trial products within the network.
   - The baseline advertisement investment is applied in the middle of each week, once a week.
   - Assume that there are 12 incandescent light bulbs having 3 hours' daily usage in each person's flat within the building, and these flats are under the same electricity tariff which is fixed at £0.20/kWh as in E.I.S. Department for Business (2021). The power consumption of the incandescent and LED light bulbs are 40W and 4.6W respectively (Du et al., 2016).
   - Assume that the price of the advertisement is £0.5 per impression for the small population network, coming from flyer and souvenir. The total budget is £110, among which £30 is investigated for the first two weeks including the free trial product.
   - Assume that the target of the product promotion programme for this small population network is to reach 70% adoption rate within 28 days.

2. Assumptions for the large population network

   - Assume that there are no free trial products in this roll out programme.
   - Assume that the electricity price is £0.20/kWh as in E.I.S. Department for Business (2021), the EEP will have 1kW power savings for 2 hours daily usage.
   - Assume that the cost of advertisement is £0.1 per thousand impressions. This figure is much smaller compared to that of the small network because for large population networks, the cost of advertisement come from online advertising which has a large volume. The total budget of advertisement is £40,000, in which £20,080 is used for advertisement in the first two weeks.

- Assume that the target adoption rate on the 28th day is 0.13%. This target value is assumed to be higher than the original adoption rate of 0.1264% obtained from the advertisement company data, which is calculated from Column Deal amount/Population in the appendix.

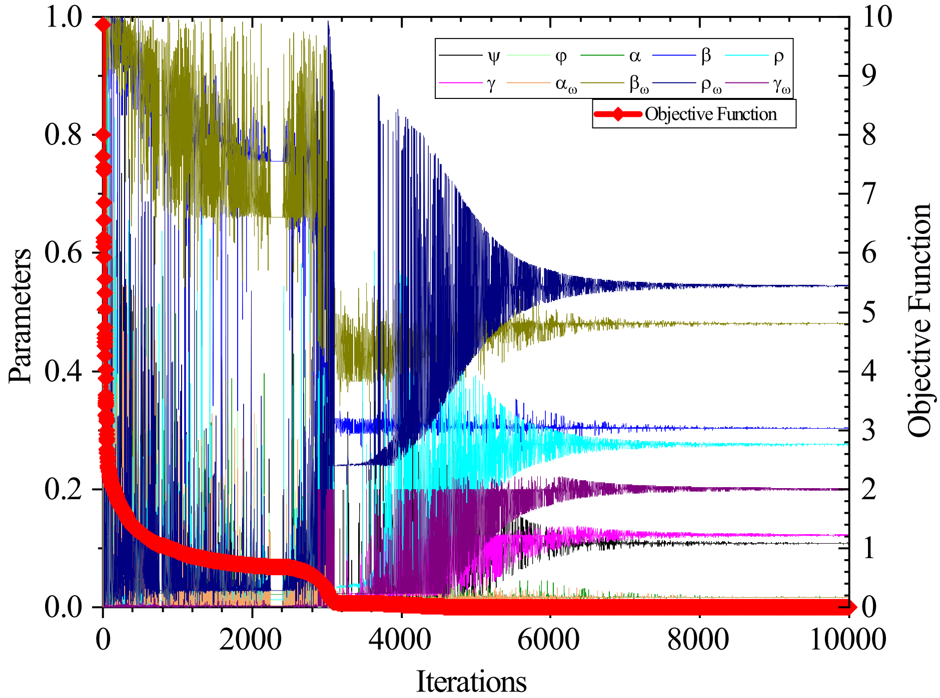


Figure 6.1: Parameter Estimation along Iterations in Advertisement Control Model (Small Network)

For both the small and the large networks, collected data on adoption rate over the first 14 days are used to determine the parameters in $\boldsymbol{\theta}$, (6.13), using the LS identification method in (6.14) and (6.15). The algorithm used for parameter estimation is genetic algorithm, the algorithm pseudo-code is similar to Algorithm 5.2 where residual function and model parameters are changed to $J_{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}$ respectively. The 'ga' function in MATLAB 2021b is chosen to perform the algorithm. The convergence graphs of the 10 parameters are shown in Fig. 6.1 for the small network, and in Fig. 6.2 for the large population network.

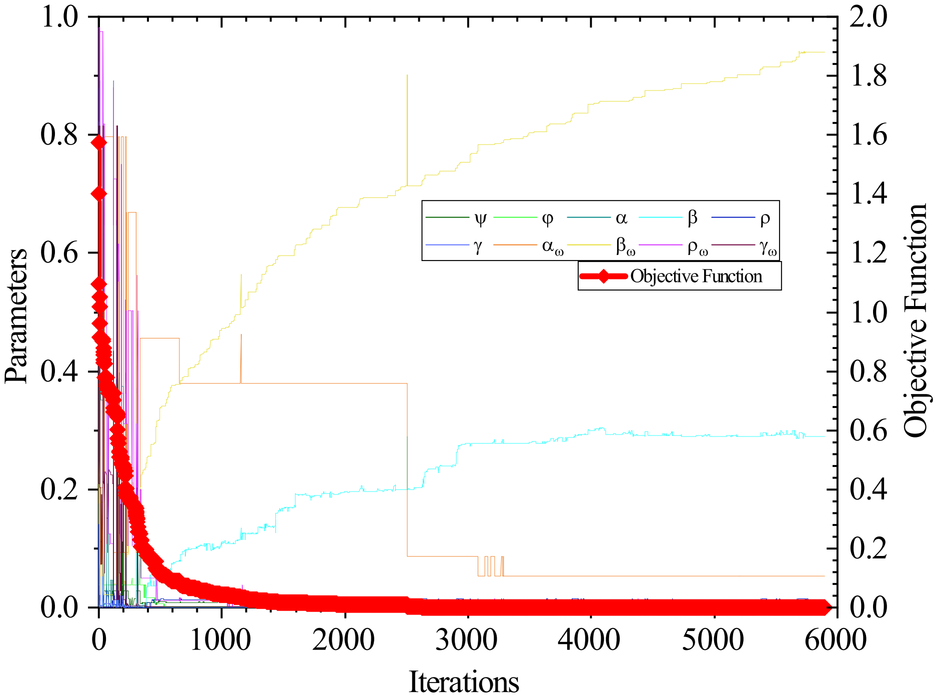


Figure 6.2: Parameter Estimation along Iterations in Advertisement Control Model (Large Network)

It can be seen from Fig. 6.1 that all parameters reach converged states at around 5,000 iterations in 'ga' algorithm. For the large network parameters estimation, the population generated in 'ga' algorithm terminates at 5,900 iterations as shown in Fig. 6.2. The objective function converges after 1,400 iterations and most parameters reach constant levels. The estimation results for the two networks are listed in Table 6.1.

Table 6.1: Estimated Model Parameters for Small and Large networks

	ψ	α	β	ρ	γ
Small network	0.1078	1.6971e-2	0.3035	0.2754	0.1218
Large network	1.0999e-2	3.8147e-6	0.2902	0.0141	4.7090e-6
	φ	α_ω	β_ω	ρ_ω	γ_ω
Small network	1.4007e-4	1.6662e-2	0.4802	0.5440	0.2000
Large network	8.1877e-6	0.0541	0.9396	3.7540e-4	5.3738e-4

6.4.2 Parameter Sensitivity Analysis

The parameter values in Table 6.1 are taken as nominal values for the model. Sensitivity analysis is applied to the parameters in a similar way to Chapter 4 and Chapter 5.

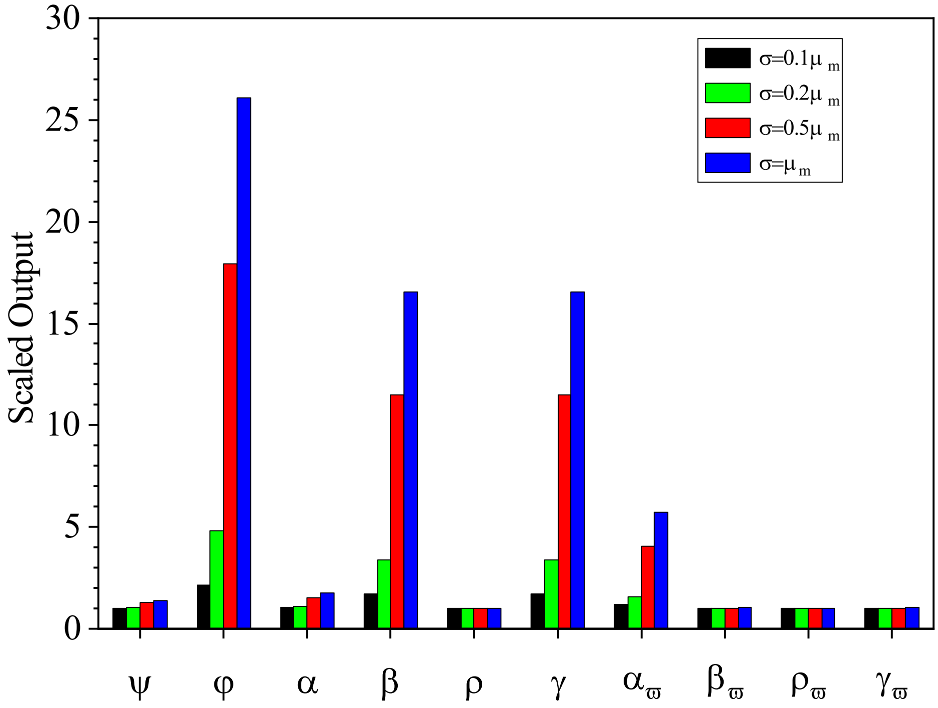


Figure 6.3: Sensitivity Analysis of Model Parameters under Four Standard Deviation Levels

Each parameter in $\boldsymbol{\theta}$ is varied in a bounded uncertainty region between zero and twice of its nominal value following a Gaussian distributions. For each parameter, nine points are uniformly sampled between the uncertainty range. The standard deviations are taken at four levels, i.e., $[0.1, 0.2, 0.5, 1]$ of the mean value. To assess the relative impacts from the ten parameters in $\boldsymbol{\theta}$, the following residual function is used as the output for the l_{sa}-th parameter:

$$J_{l_{sa}} = \sum_{\tau=1}^{M+K} \sum_{l=1}^{9} \Big(R\big(\tau, \boldsymbol{\theta}_{l_{sa}}(l)\big) - y_R(\tau) \Big)^2 \cdot \boldsymbol{pdf}_{l_{sa}}(l) \tag{6.29}$$

where $R\big(\tau, \boldsymbol{\theta}_{l_{sa}}(l)\big)$ is the adoption rate calculated from the model, $y_R(\tau)$ is the adoption rate obtained from the survey data at time τ, $\boldsymbol{pdf}_{l_{sa}}(l)$ is the pdf of the $l-$th sampled point for the l_{sa}-th parameter.

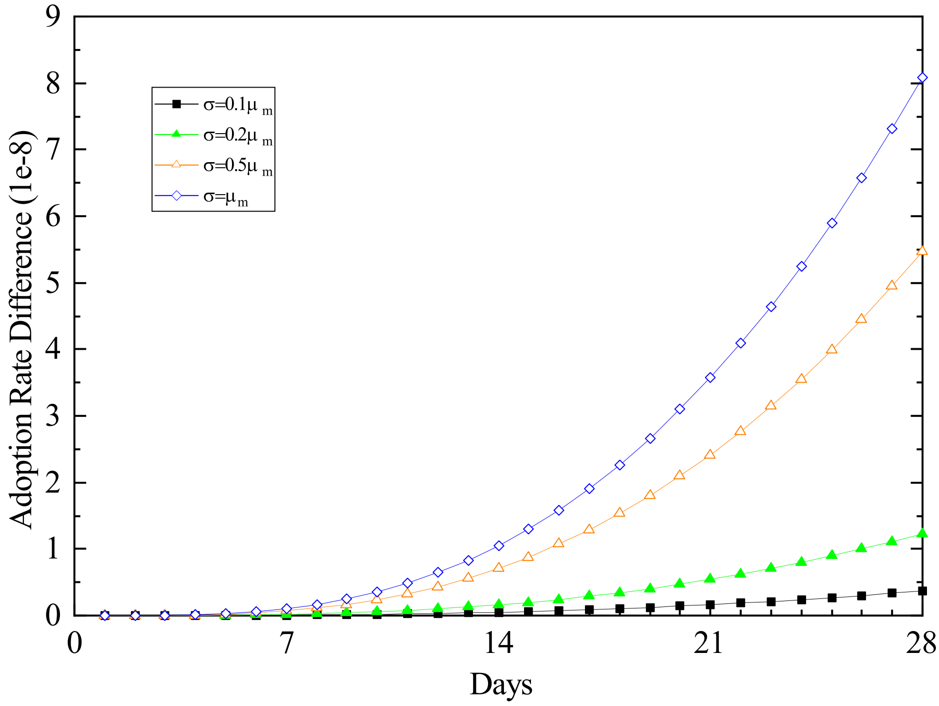


Figure 6.4: Scaled Adoption Rate Difference Between the Nominal Value and Monte-Carlo Calculation Under Parameter Variation (Parameter φ, Standard Deviations: 10%, 20%, 50% and 100% of Mean Value)

The results of $J_{l_{sa}}$ are used for sensitivity analysis. Taking the standard deviation at four levels, i.e., 10%, 20%, 50% and 100% of the nominal value, the sensitivity graph is shown in Fig. 6.3, where the output $J_{l_{sa}}$ is normalized by the value calculated at the nominal parameters for $\boldsymbol{\theta}_{l_{sa}}$.

From the sensitivity analysis results in Fig. 6.3, it can be observed that among the ten parameters, φ, β, γ, and α_ω have stronger influence on the formed output function in (6.29), φ is the most sensitive parameter. Next, this parameter is chosen for further calculations under four different standard deviations. Again, the same sampling of nine points is applied to φ, and the calculated adoption rates are weighted by pdf values corresponding to the sampled value of φ, i.e., $R_\varphi(\tau) = \sum_{j=1}^{9} R\big(\tau, \varphi(l)\big) \cdot \boldsymbol{pdf}_\varphi(l)$.

The daily adoption rates calculated by this Monte-Carlo study is compared against the results by nominal value. The difference between them are shown in Fig. 6.4, from which it can be seen that when the parameter uncertainty is larger (larger standard deviation), the difference between the model result and the surveyed data is larger as expected. However, the difference caused by model uncertainty seems to be very small, at a level of $1e-8$, over the 28 days. This indicates the robustness of the proposed model against parameter variations to some extent.

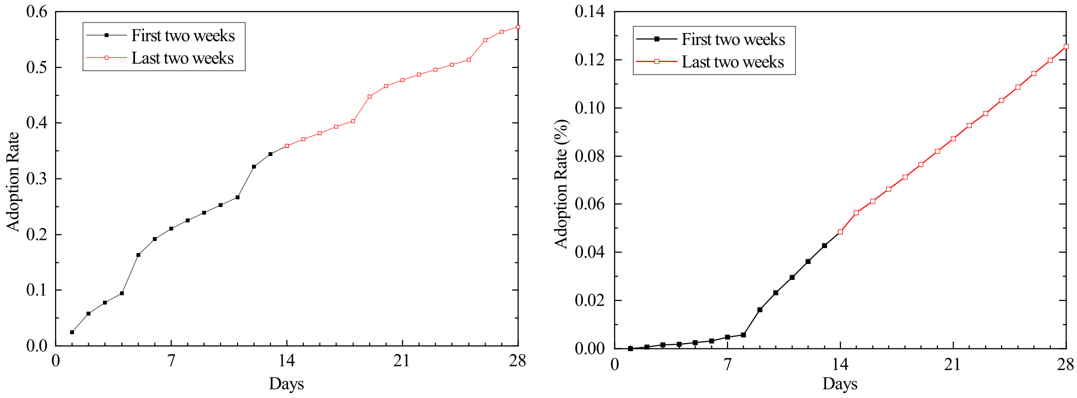
6.4.3 Advertisement Control Design for Small Population Network

The advertisement control strategies are applied to the small population network. The target EEP adoption rate is set to be 70% for the small network, the budget limit is £110.

6.4.3.1 Baseline Strategy

At the beginning of the programme, the initial values for the three states in model (6.9) are set to be $S(0) = 0.925$, $I(0) = 0$, and $R(0) = 0.025$ by (6.17) with $z = 1, N = 40$. It should be mentioned that two individuals are excluded in this social network as they show no interest on the energy saving product during the survey. The time period is 28 days and the sampling rate is one day for each time point. The impressions are exerted on four days only, i.e., Day 4, Day 11, Day 18 and Day 25, with the impression magnitude of 20 for each day. For the other 24 days, the impressions are zero under the baseline strategy.

The adoption rate is calculated from the information diffusion model without advertisement input as in (6.6). The results are shown in Fig. 6.5a, in which the EEP adoption rate values over the first two weeks are represented with solid square marker in black, and the next two weeks with red square hollow markers. It can be seen that the EEP adoption rate is 35.88% at the end of Week 2, and is 57.24% at the end of Week 4. The Day-28 adoption rate is lower than 60%, and the target level is 70%. Therefore, more advertisement input is required in order to accomplish the target over the period of Day 15 to Day 28. The costs of advertisement during the first 14 days and last 14 days are calculated by (6.19) to be the same as £20 for each 14 days, the total advertisement cost is £40 for the baseline strategy. In the next section, the four proposed optimization design methods are applied to achieve the target adoption rate.



(a) Small Network (b) Large Network

Figure 6.5: Baseline Daily Adoption Rates for Small and Large Networks (the First 14 days Data from Survey, the Last 14 days Data from Model Calculation under Baseline Advertisement Impressions)

6.4.3.2 Method 1: Targeting for Lowest Advertisement Cost

The first design method is to achieve the target EEP adoption rate with the smallest advertisement cost, as formulated in (6.23). The design results for optimization in (6.25) are listed in Table 6.2 covering the daily impressions from Day 15 to Day 28. The convergence graph is shown in Fig. 6.6. The results of daily impressions from other design methods are also listed in this table for comparison purpose.

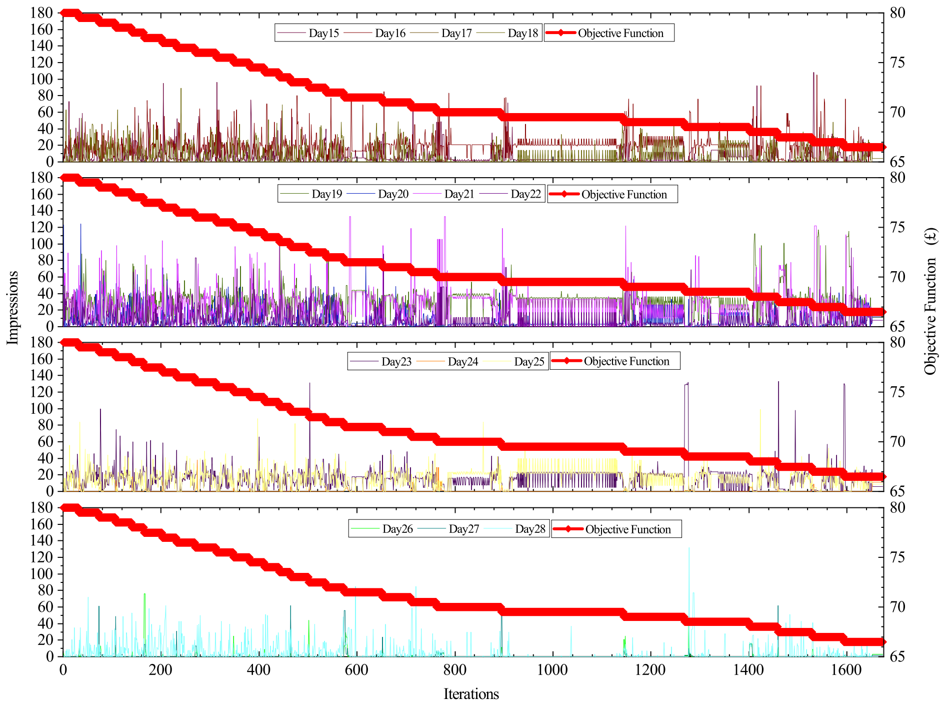


Figure 6.6: Convergence Graph of Impression Optimisation Using Advertisement Control for Small Population Network with Method 1

The total cost on advertisement from Day 15 to Day 28 can be calculated by (6.19) to be £66.5($= 0.5 * 133$). Considering the cost in the first 14 days using the baseline strategy (£30), the total advertisement cost over the four weeks is £96.5, which is within the budget limit of £110. It can be seen that the number of daily impressions is between 2 to 18 from Day 15 to Day 27, which are displayed as the green colour curve in Fig. 6.7.

Table 6.2: Optimal Advertisement Investment of Different Methods in Small Population Network

Impressions\Days	15	16	17	18	19	20	21
Baseline	0	0	0	20	0	0	0
Method 1	17	18	17	4	11	12	15
Method 2	25	22	20	16	21	16	10
Method 3	61	37	29	21	12	0	0
Method 4	17	14	13	13	13	13	14
Impressions\Days	22	23	24	25	26	27	28
Baseline	0	0	0	20	0	0	0
Method 1	8	6	11	9	3	2	0
Method 2	24	6	0	0	0	0	0
Method 3	0	0	0	0	0	0	0
Method 4	15	13	13	13	9	0	0
Total	X_T	C_{adv} (£)	C_{total} (£)		$R(28)$		
Baseline	**40**	**20**	**50**		**0.5724**		
Method 1	**<u>133</u>**	**<u>66.5</u>**	**<u>96.5</u>**		**0.7002**		
Method 2	**160**	**80**	**110**		**0.7222**		
Method 3	**160**	**80**	**110**		**0.7048**		
Method 4	**160**	**80**	**110**		**0.7272**		

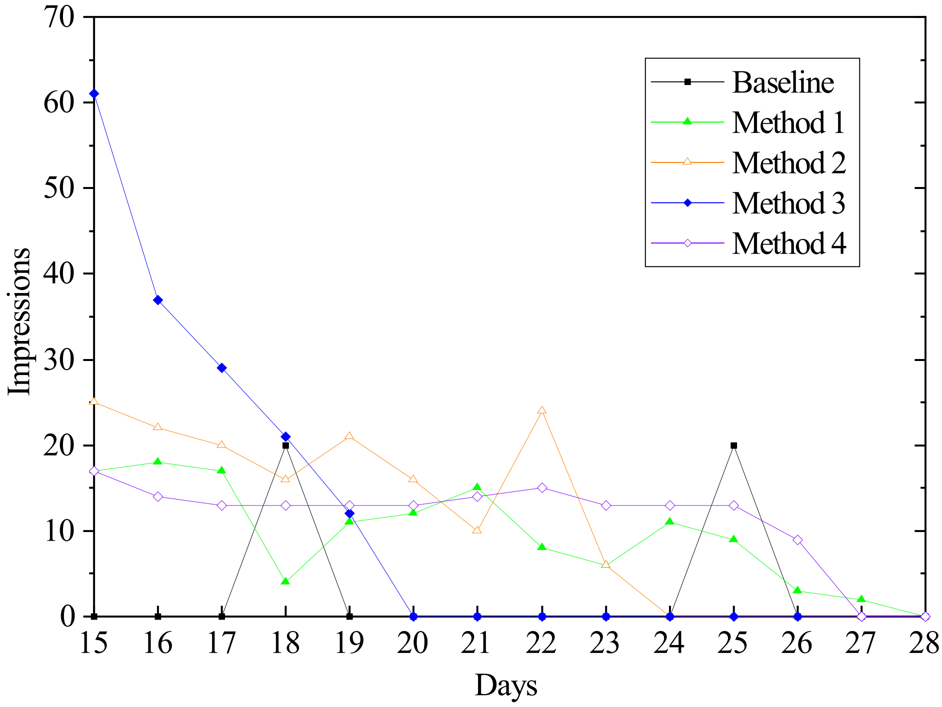


Figure 6.7: Daily Impressions under Different Strategies in Small Population Network

Next the daily adoption rate is calculated and compared with the baseline strategy, as illustrated in Fig. 6.8, the green curve for Method 1. The numerical figures are provided in Table 6.3 for a closer check. It can be seen that the adoption rate under Method 1 reaches 70.02% at the end of Week 4 (Day 28), which is above the the target rate of 70%.

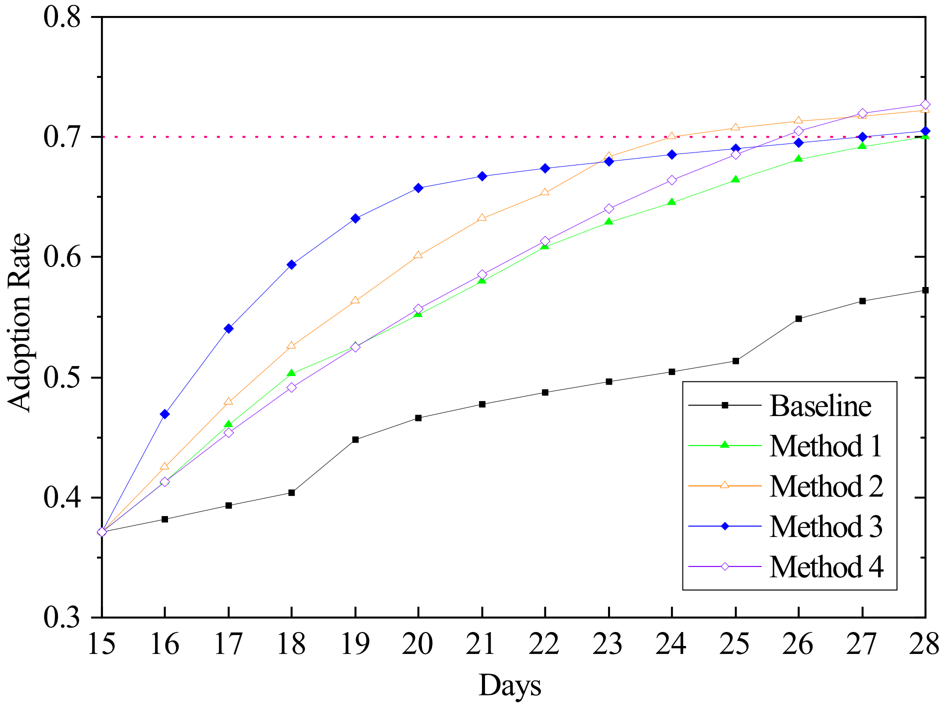


Figure 6.8: Adoption Rates under Different Programmes of Small Network (One Baseline Strategy and Four Design Methods, Target Level at 0.7)

Table 6.3: Daily Adoption Rate of Different Methods in Small Population Network

Adoption rate/Days	15	16	17	18	19	20	21
Baseline	0.3708	0.3821	0.3930	0.4036	0.4479	0.4660	0.4775
Method 1	0.3708	0.4131	0.4600	0.5031	0.5258	0.5524	0.5796
Method 2	0.3708	0.4256	0.4793	0.5259	0.5632	0.6012	0.6319
Method 3	0.3708	0.4690	0.5409	0.5939	0.6321	0.6572	0.6676
Method 4	0.3708	0.4131	0.4542	0.4911	0.5251	0.5566	0.5856
Adoption rate/Days	**22**	**23**	**24**	**25**	**26**	**27**	**28**
Baseline	0.4872	0.4962	0.5049	0.5134	0.5487	0.5632	0.5724
Method 1	0.6086	0.6292	0.6452	0.6642	0.6813	0.6920	**<u>0.7002</u>**
Method 2	0.6537	0.6839	**<u>0.7000</u>**	0.7076	0.7130	0.7177	0.7222
Method 3	0.6743	0.6799	0.6852	0.6902	0.6952	**<u>0.7001</u>**	0.7048
Method 4	0.6136	0.6406	0.6639	0.6852	**<u>0.7048</u>**	0.7201	0.7272

6.4.3.3 Method 2: Targeting for Shortest Time

Figure 6.9: Convergence Graph of Impression Optimisation Using Advertisement Control for Small Population Network with Method 2

The second design method is to reach the target EEP adoption rate within the shortest time, as in (6.26). The optimization convergence results are shown in Fig. 6.9. The numerical results of daily impressions from Day 15 to Day 28 are listed in Table 6.2. The shortest time to reach the target adoption rate is found to be nine days from Day 15 to Day 23.

The higher investment on advertisement in the early period leads to a fast-growing curve of the adoption rate (see Fig. 6.8, line with orange colour hollow triangles) and reaches the target of 70% on Day 24 at 70.00% (see Table 6.3), 4 days ahead of the end of the programme. With this method, the total cost on advertisement (£110) is all used, which is a price paid to achieve the target rate within a shorter time. In this case, the overall cost benefit can be expected from the saved energy cost. It should be noted that the adoption rate after Day 24 continues to have small increase although there're no more advertisement exerted in the last four days. This minor increase is due to the information diffusion mechanism of the network itself.

6.4.3.4 Method 3: Targeting For Largest Energy Savings

The advertisement input is designed to achieve the largest energy savings, as in (6.27). The convergence graph of the optimization process is shown in Fig. 6.10. The daily impressions under Method 3 are listed in Table 6.2, which shows that large amount of advertisement investment is made at the beginning of Week 3. This suggests that the total energy savings are highly related to early-stage adoption rate. The daily impressions in Method 3 are shown in Fig. 6.7 (blue line), in which the largest figure is on the first day (Day 15) followed by gradually decreased daily impressions in the next four days, and then becomes zeros from Day 20. This decrease of daily impressions is due to the diminishing return effect, as shown from the BFT update of the adoption rate and the positive response rate in (6.11) and (6.12). The daily adoption rate values under Method 3 are shown in Fig. 6.8 with blue solid diamond markers. By applying this method, again all the budget is required to achieve the largest energy savings.

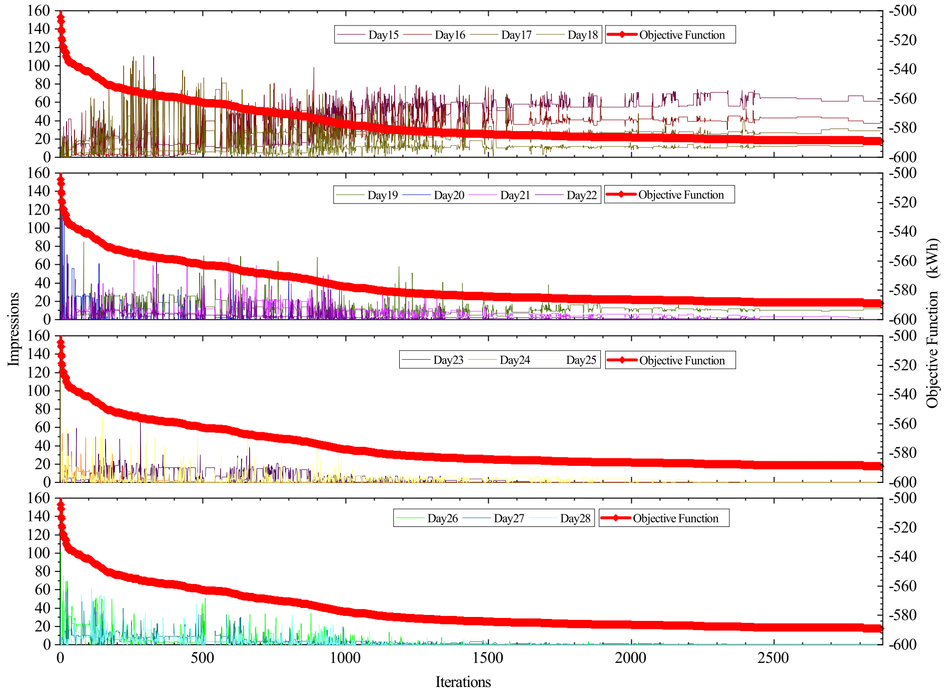


Figure 6.10: Convergence Graph of Impression Optimisation Using Advertisement Control Model for Small Population Network with Method 3 (28 days)

The electricity savings of all methods for the last 14 days are calculated by (6.16) and shown in Fig. 6.11 with a comparison made with the baseline strategy. The energy savings achieved by one adoption is $E = 1.2744$ kWh as explained in Section 6.4.1. This comparison shows that among all methods, Method 3 has achieved the largest energy savings. The total energy savings over the last 14-day period is 446.6 kWh with Method 3 and 337.8 kWh with the baseline plan that are shown in Table 6.5 respectively. There's an increase of 32.21% on energy savings with Method 3 compared to the baseline result. This method shows the significance of early adoption of EEP in total energy savings.

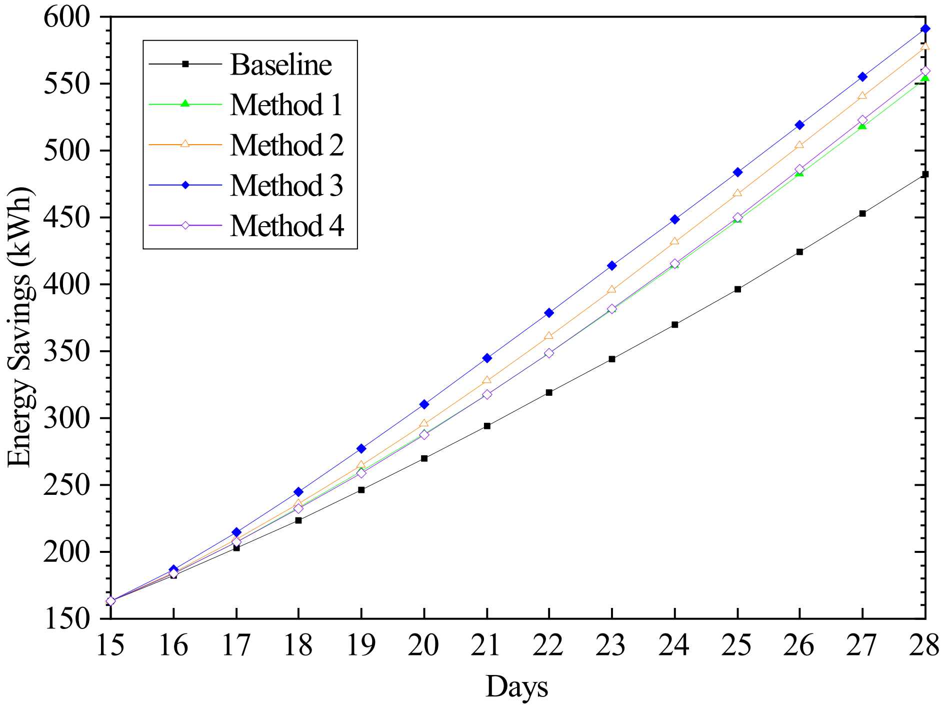


Figure 6.11: Daily Energy Savings of Small Network (Baseline Strategy and Four Design Methods)

Table 6.4: EES of Baseline Method for First 14 Days in Small Population Network

EES (kWh)/Days	1	2	3	4	5	6	7
Baseline	1.3	4.2	8.2	13.0	21.3	31.1	41.8
EES (kWh)/Days	8	9	10	11	12	13	14
Baseline	53.3	65.5	78.4	92.0	108.4	126.0	144.2

Using the assumed electricity tariff in Section 6.4.1, $\pi = 0.20$ £/kWh, the electricity bill reduced for the network is calculated as £28.84 over the first 14 days with EES from Table 6.5, £89.32 over the last 14 days with Method 3 from Table 6.5. Thus the energy savings benefit for the whole 28-day period with Method 3 is £118.16 which is larger than the total budget (£110) spent on using advertisement for product promotion.

Table 6.5: EES of Different Methods for Last 14 Days in Small Population Network (Cumulate from Day 15)

EES (kWh)/Days	15	16	17	18	19	20	21
Baseline	18.9	38.4	58.4	79.0	101.8	125.6	149.9
Method 1	18.9	40.0	63.4	89.1	115.9	144.0	173.6
Method 2	18.9	40.6	65.0	91.8	120.5	151.2	183.4
Method 3	18.9	42.8	70.4	100.7	132.9	166.4	200.4
Method 4	18.9	40.0	63.1	88.1	114.9	143.3	173.1
EES(kWh)/Days	22	23	24	25	26	27	28
Baseline	174.7	200.0	225.8	251.9	279.9	308.6	337.8
Method 1	204.6	236.7	269.6	303.4	338.1	373.4	409.1
Method 2	216.7	251.6	287.3	323.3	359.7	396.3	433.1
Method 3	234.8	269.4	304.4	339.6	375.0	410.7	446.6
Method 4	204.4	237.1	270.9	305.8	341.8	378.5	415.5

6.4.3.5 Method 4: Targeting for Largest Installation Number

The fourth design method is applied to achieve the largest installation number, or equivalently the largest adoption rate. The convergence graph of the optimization process is shown in Fig. 6.12. The results of daily impressions under Method 4 are listed in Table 6.2, showing that the advertisement investment from Day 15 to Day 25 follows a relatively smooth pattern. The daily adoption rate profile of Method 4 is shown in Fig. 6.8 by the violet hollow diamond markers. It can be seen from Table 6.3 that the adoption rate reaches 72.72% on Day 28, the largest among the four design methods. Comparing to the baseline strategy (57.24%), the strategy in Method 4 has an increase of 15.48% in adoption rate.

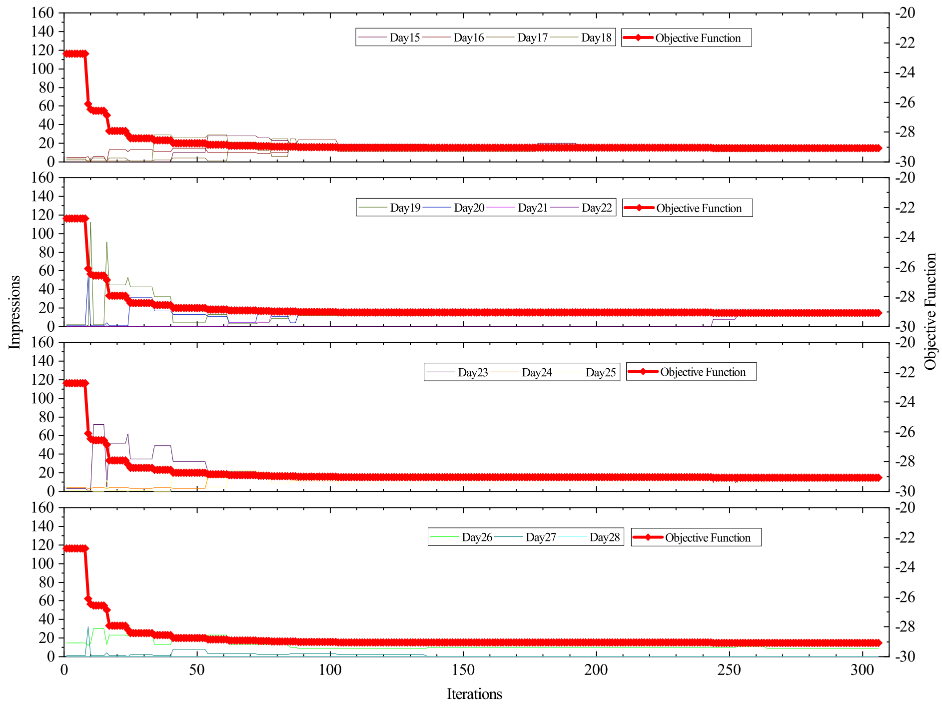


Figure 6.12: Convergence Graph of Impression Optimisation Using Advertisement Control for Small Population Network with Method 4

6.4.3.6 Cost Effectiveness Analysis

The cost savings can be calculated using electricity tariff π and EES_T, (6.20). The results are shown in Fig. 6.13. Taking the cost savings as 'profit' while the advertisement and initial free product as 'cost', the payback period can be calculated by analysing the daily balance which is shown in Fig. 6.14. It can be seen that Method 1 has the shortest payback period which is on Day 26. Method 3 has the largest daily deficit of £55.02, among all four methods, which is shown in Table 6.6.

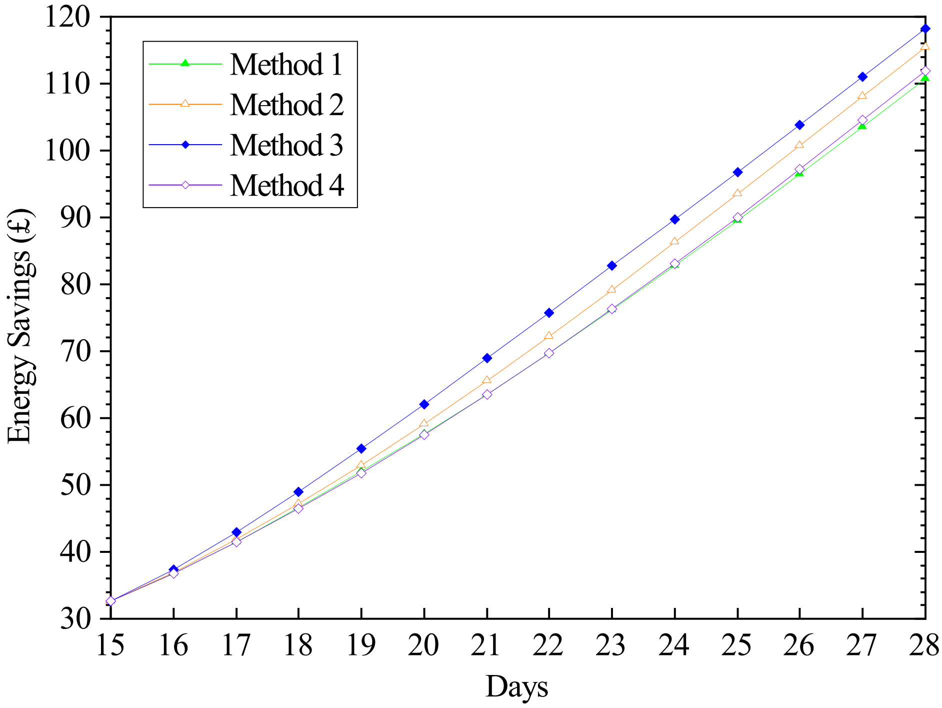


Figure 6.13: Cost Savings of Small Network (Four Design Methods)

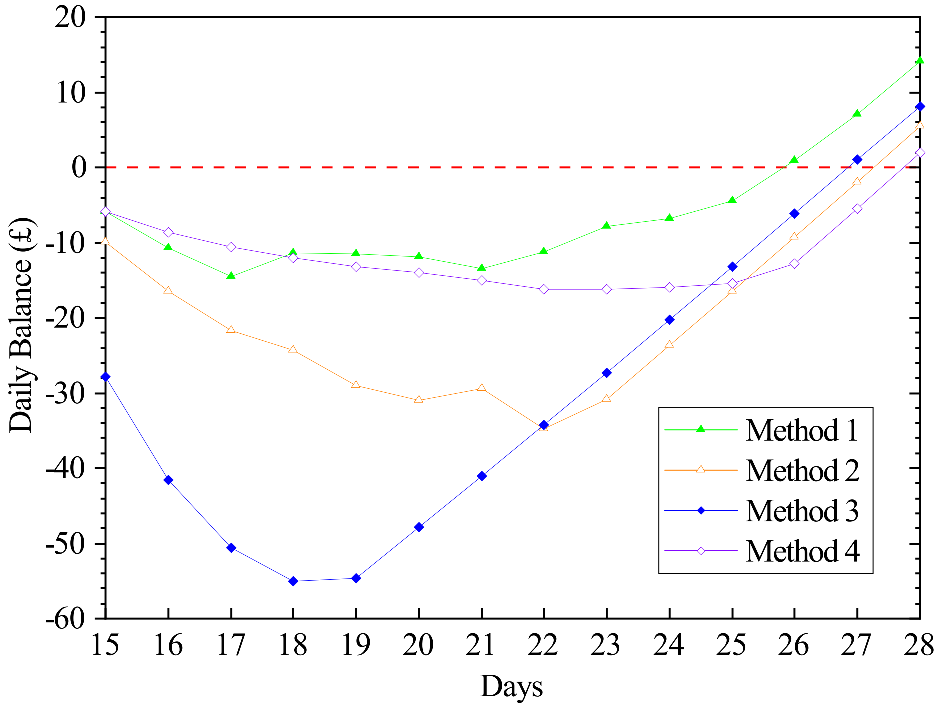


Figure 6.14: Daily Deficit/Surplus of Small Network (Four Design Methods)

Table 6.6: Daily Balance of Different Methods in Small Population Network

Daily balance(£)/Days	15	16	17	18	19	20	21
Method 1	-5.87	-10.66	-14.47	-11.34	-11.48	-11.85	-13.44
Method 2	-9.87	-16.54	-21.65	-24.29	-29.05	-30.92	-29.47
Method 3	-27.87	-41.59	-50.58	-55.02	-54.58	-47.88	-41.07
Method 4	-5.87	-8.66	-10.53	-12.03	-13.17	-14.00	-15.03
Daily balance(£)/Days	22	23	24	25	26	27	28
Method 1	-11.24	-7.82	-6.75	-4.47	0.97	7.03	14.17
Method 2	-34.81	-30.84	-23.70	-16.49	-9.22	-1.90	5.46
Method 3	-34.20	-27.27	-20.28	-13.24	-6.16	0.98	8.17
Method 4	-16.27	-16.24	-15.97	-15.49	-12.80	-5.46	1.90

6.4.3.7 Summary for Small Network Simulation

For a small population network, the information diffusion within the network contributes a lot to the adoption rate achieved. This can be seen from Fig. 6.8 that even under zero impressions for some days, e.g., the last few days under Method 2 and Method 3, the adoption rate still continues to grow although gradually. Meanwhile, advertisement also has a strong impact on the adoption rate, all advertisement control methods achieve a clear increase in the final adoption rates compared to the baseline strategy. For Method 3, the adoption rate has an increase from 37.08% to 65.72% within only five days (Day 15 to Day 20) (Table 6.3), and the EES for the last 14 days increases 32.21% compared to the baseline strategy Table 6.5. It can also be seen from Fig. 6.11 that Method 3 achieves the largest energy savings among the four methods as it has the objective to maximise the total energy savings.

6.4.4 Advertisement Control of Large Population Network

6.4.4.1 Baseline Strategy

The collected data from an advertisement company for a large population network with one million users is attached in the appendix. The target adoption rate is 0.13%. The period of study is 28 days.

Table 6.7: Original Impression Data in Large Population Network (from Survey)

Day	1	2	3	4	5	6	7
Impressions (1e3)	7,247	7,316	6,506	4,250	4,157	2,768	7,798
Day	8	9	10	11	12	13	14
Impressions (1e3)	7,358	64,900	12,851	32,333	21,386	10,676	11,258
Day	15	16	17	18	19	20	21
Impressions (1e3)	8,626	11,158	10,989	10,986	12,283	9,651	10,606
Day	22	23	24	25	26	27	28
Impressions (1e3)	8,587	9,531	10,508	12,027	9,030	10,303	10,315

From the survey data shown in Table 6.7, it can be seen that a total amount of £20,080 is used for advertisement in the first two weeks, and £14,460 for the next two weeks. The adoption rate is calculated to be 0.0486% at the end of Week 2 from the survey data.

Based on the developed model, the adoption rates are calculated for the next 14 days using the baseline strategy, which gives the adoption rate of 0.1254% on Day 28. This figure is lower than the target level of 0.13% set up for optimization design. The adoption rates in the first 14 days (surveyed) and the last 14 days (modelled) are represented by black solid diamonds and red hollow diamonds in Fig. 6.5b. Similar to the small population network case study, four advertisement control methods are applied to achieve the target adoption rate under the given budget of £40,000 over 28 days.

6.4.4.2 Method 1: Targeting for Lowest Advertisement Cost

With the optimization design in (6.25) to minimize the advertisement cost, the daily impressions are given in Table 6.8. The results of daily impressions from other design methods are also listed in the same table for comparison. It can be seen that the total number of impressions for Method 1 is the lowest among the four design methods, which means the advertisement cost is the smallest with Method 1. The convergence graph of the optimization process is shown in Fig. 6.15.

Table 6.8: Optimal Advertisement Investment of Different Methods in Large Population Network

Impressions (1e3) \Days	15	16	17	18	19	20	21
Method 1	155,598	37	1,762	341	260	1,742	408
Method 2	1	0	3	374	1	0	0
Method 3	192,935	3,153	372	93	303	27	1,914
Method 4	198,844	277	79	0	0	0	0
Impressions (1e3) \Days	22	23	24	25	26	27	28
Method 1	69	306	1,169	2,668	38	366	0
Method 2	0	1	198,470	0	0	1	0
Method 3	358	7	21	4	11	2	0
Method 4	0	0	0	0	0	0	0
Total	X_T	C_{adv}(£)	C_{total}(£)		$R(28)(\%)$		
Method 1	**164,664**	**16,466**	**36,546**		**0.1300**		
Method 2	**198,851**	**19,885**	**39,965**		**0.1342**		
Method 3	**199,200**	**19,920**	**40,000**		**0.1349**		
Method 4	**199,200**	**19,920**	**40,000**		**0.1349**		

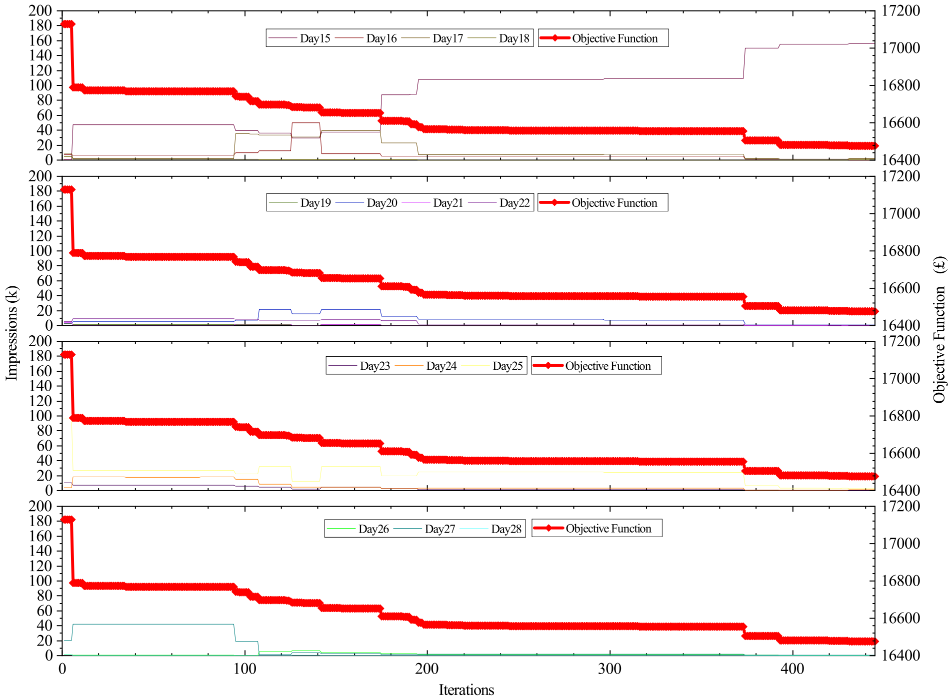


Figure 6.15: Convergence Graph of Impression Optimisation Using Advertisement Control or Large Population Network with Method 1

Figure 6.16 shows a comparison of the daily impressions between the four control methods and the baseline strategy, from which it can be seen that the advertisement investment is mostly spent in Week 3 for Method 1, to be more specific, on the first day (Day 15) of that week.

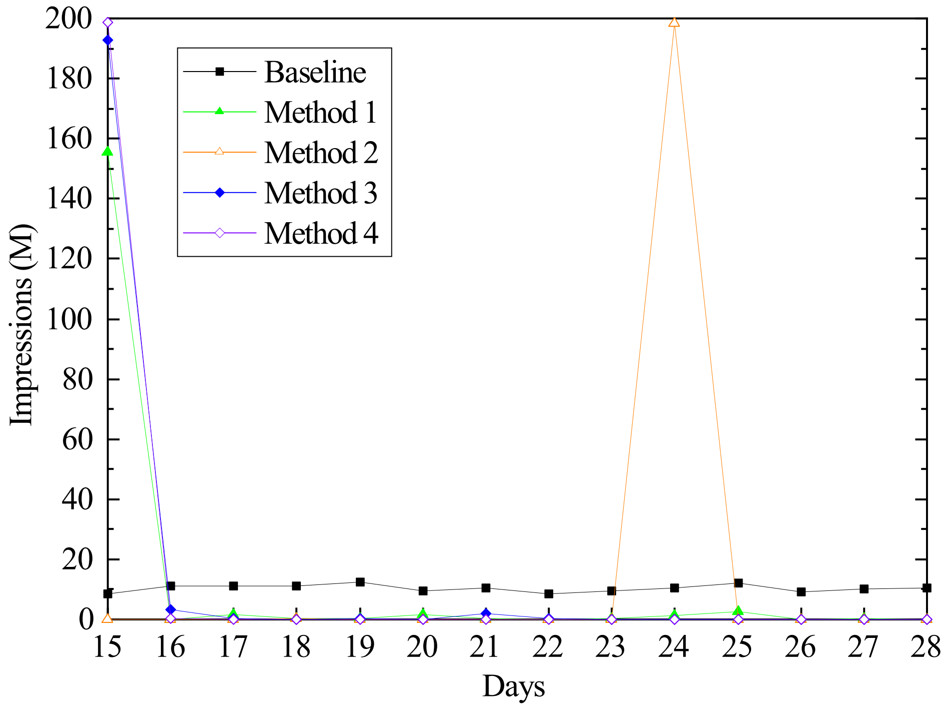


Figure 6.16: Daily Impressions for Different Strategies in Large Population Network

The investment cost in Method 1 is calculated to be £36, 546.4 which is within the budget. The adoption rates from the optimal solution and the baseline result are shown in Fig. 6.17. The baseline forecast results are represented in black curves. The optimal solution results under the four methods are shown in green, orange, blue and violet curves, respectively. As shown in Table 6.9, the optimal solution in Method 1 reaches the target adoption rate of 0.13% on Day 28.

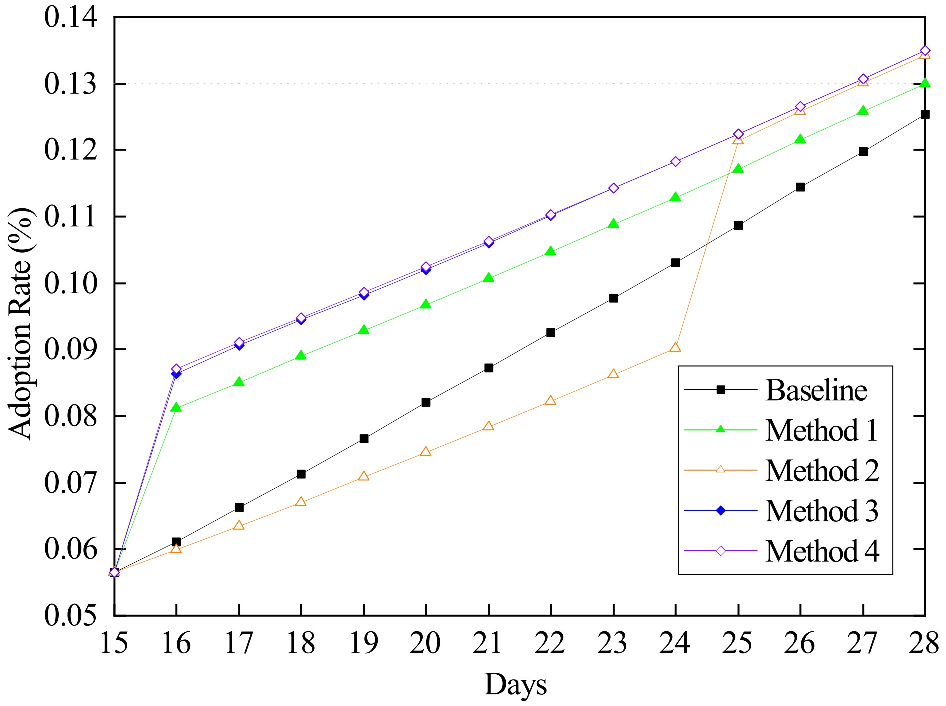


Figure 6.17: Adoption Rates of Four Designs for Large Population Network (Four Optimal Design Results and the Baseline Result)

Table 6.9: Daily Adoption Rate of Different Methods in Large Population Network

Adoption rate(%)/Days	15	16	17	18	19	20	21
Baseline	0.0565	0.0611	0.0662	0.0713	0.0765	0.0820	0.0872
Method 1	0.0565	0.0812	0.0851	0.0890	0.0928	0.0966	0.1007
Method 2	0.0565	0.0600	0.0635	0.0671	0.0708	0.0745	0.0784
Method 3	0.0565	0.0863	0.0907	0.0944	0.0982	0.1021	0.1060
Method 4	0.0565	0.0871	0.0911	0.0948	0.0986	0.1024	0.1063
Adoption rate(%)/Days	**22**	**23**	**24**	**25**	**26**	**27**	**28**
Baseline	0.0926	0.0977	0.1031	0.1086	0.1144	0.1198	0.1254
Method 1	0.1047	0.1088	0.1128	0.1171	0.1216	0.1258	**0.1300**
Method 2	0.0823	0.0862	0.0902	0.1214	0.1259	**0.1300**	0.1342
Method 3	0.1102	0.1143	0.1183	0.1224	0.1265	**0.1307**	0.1349
Method 4	0.1103	0.1143	0.1183	0.1224	0.1265	**0.1307**	0.1349

6.4.4.3 Method 2: Targeting for Shortest Time

The optimal control solution on daily impressions can be seen in Table 6.8. The optimization convergence graph is shown in Fig. 6.18. The investment in this solution is £39, 965, which is also within the budget limit, and the adoption rate is shown in Fig. 6.17 by the orange curve. The adoption rate reaches 0.1307% on Day 27 (see Table 6.9), which is one day ahead of the end of the programme.

Figure 6.18: Convergence Graph of Impression Optimisation Using Advertisement Control for Large Population Network with Method 2

6.4.4.4 Method 3: Maximize for Total Energy Savings

The energy savings achieved by one adoption is $E = 2$ kWh as explained in Section 6.4.1. Using Method 3, the calculated daily impressions are listed in Table 6.8. The optimization convergence graph is shown in Fig. 6.19. The adoption rate is shown by the blue curve in Fig. 6.17. There is a high number of impressions on the first day of week 3 (Day 15). This method uses all £40, 000 budget to achieve the designed target with an adoption rate of 0.1349% achieved on Day 28 which is shown in Table 6.9.

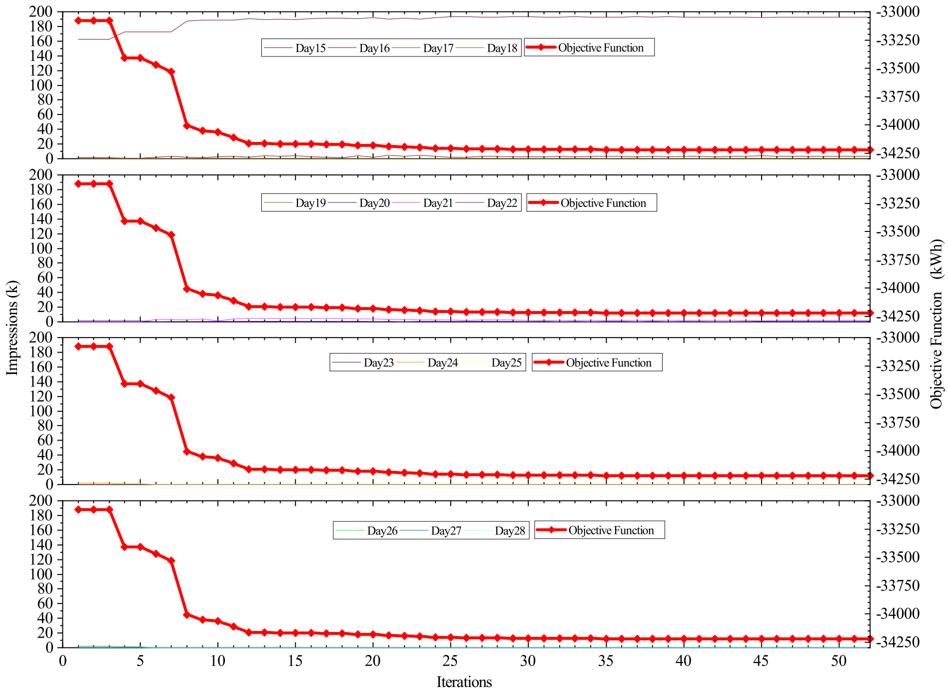


Figure 6.19: Convergence Graph of Impression Optimisation Using Advertisement Control for Large Population Network with Method 3

The energy savings profiles of all methods are shown in Fig. 6.20, where the baseline strategy data and the optimal control results are represented. The optimal solution of Method 3 brings a total saving of 34.223 MWh over 28 days which is shown in Table 6.10. For the last 14 days, Method 3 has 29.887 MWh energy savings, while the baseline strategy has 25.249 MWh. This indicates an increase of 18.37% in energy savings due to the advertisement control strategy. Using the assumed electricity tariff in Section 6.4.1, the electricity bill reduced for the network over the whole 28-day period with Method 3 is £7,049.8 which is smaller than the total budget (£40,000) that is spent using advertisement for product promotion.

Table 6.10: Cumulated EES of Different Methods in Large Population Network

EES(kWh)/Days	1	2	3	4	5	6	7
Baseline	0	12	42	80	132	194	288
EES(kWh)/Days	8	9	10	11	12	13	14
Baseline	402	726	1,192	1,786	2,510	3,364	4,336
EES(kWh)/Days	15	16	17	18	19	20	21
Baseline	5,467	6,690	8,013	9,440	10,970	12,610	14,354
Method 1	5,467	7,091	8,793	10,573	12,428	14,361	16,376
Method 2	5,467	6,666	7,935	9,277	10,692	12,183	13,751
Method 3	5,467	7,209	9,031	10,928	12,900	14,948	17,074
Method 4	5,467	7,193	9,007	10,896	12,860	14,901	17,021
EES(kWh)/Days	22	23	24	25	26	27	28
Baseline	16,205	18,160	20,222	22,394	24,681	27,077	29,585
Method 1	18,471	20,646	22,903	25,245	27,676	30,191	32,791
Method 2	15,396	17,121	18,925	21,353	23,870	26,471	29,156
Method 3	19,279	21,565	23,931	26,379	28,910	31,524	34,223
Method 4	19,225	21,510	23,876	26,324	28,855	31,469	34,167

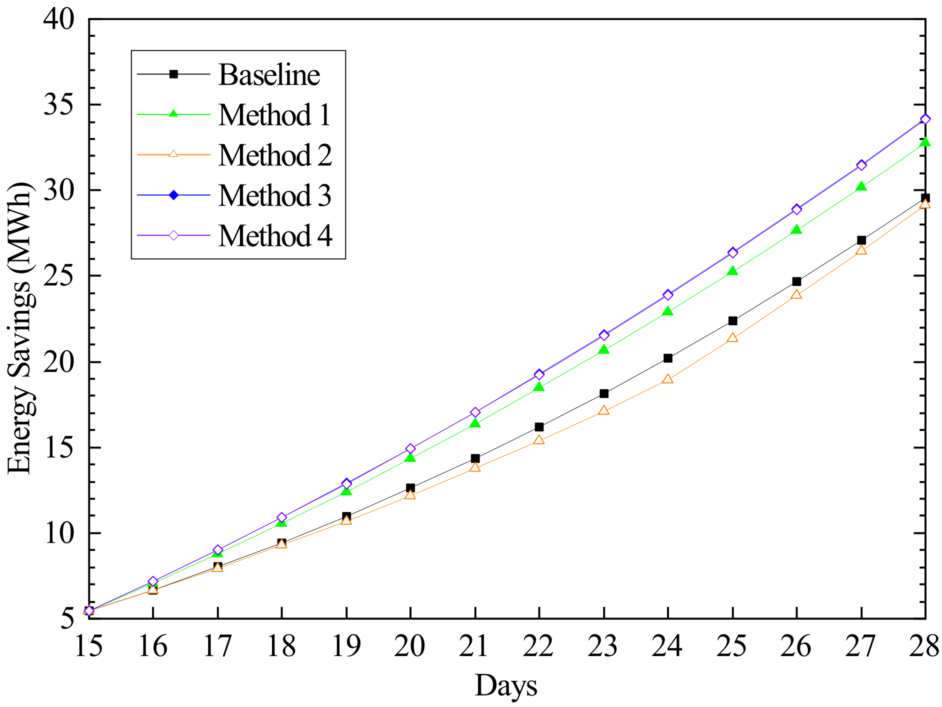


Figure 6.20: Daily Energy Savings of Large Network (Baseline Strategy and Four Design Methods)

6.4.4.5 Method 4: Maximize the Adoption Rate

It is observed from Table 4 that the optimal solution on the daily impressions of Method 4 is similar to Method 3. The convergence graph is shown in Fig. 6.21. The adoption rate of Method 4 is shown in Fig. 6.17 with violet dashed curve, and the adoption rate reaches 0.1349% on Day 28, which is above the programme target of 0.13%. The cost of this method is the £40, 000, the total budget allowed. The number of product installations for the whole network is 1, 349. With this design, most of the advertisement investment is used for the first day of the design period (Day 15), which is likely due to the diminishing return effect.

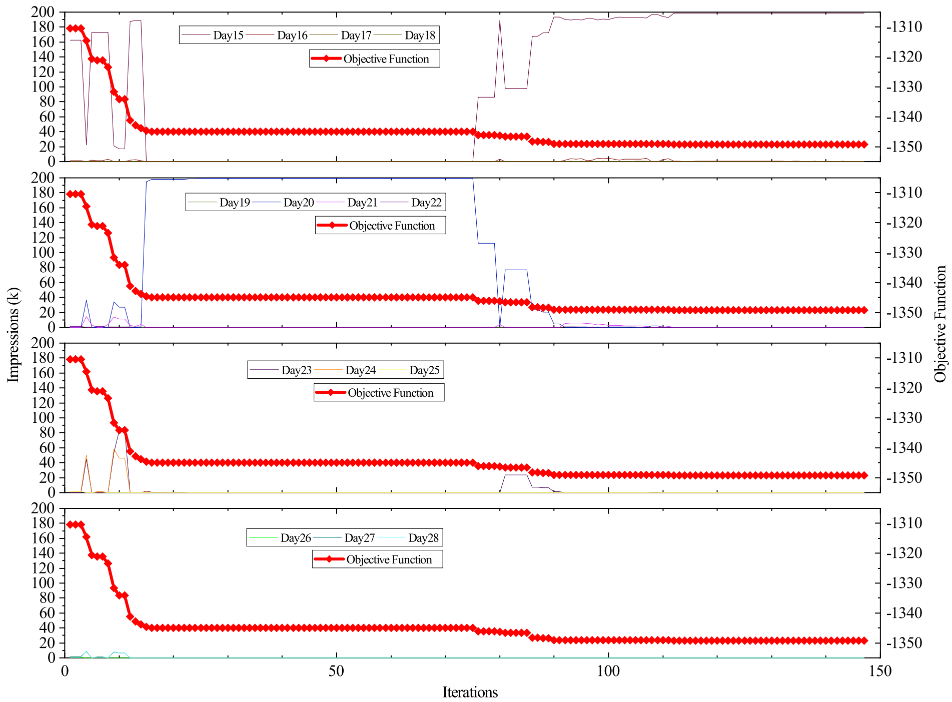


Figure 6.21: Convergence Graph of Impression Optimisation Using Advertisement Control for Large Population Network with Method 4

For large population networks, the advertisement investment seems to have more effective influence on the adoption rates compared to the small population network. It can be seen from Fig. 6.17 that the adoption rates are significantly increased through advertisement control, and Method 3 achieves the largest total energy savings among the four design methods (see Fig. 6.20).

6.4.4.6 Cost Effectiveness Analysis

The cost savings of large network can be calculated using electricity tariff π and EES_T which is shown in Fig. 6.22. It can be seen that even with Method 3 that has the highest daily energy savings, the cost saving within 28 days is less than the advertisement investment. Taking the energy cost savings as 'profit' and the advertisement expense as 'cost', the payback period of the large network must be larger than 28 days for all the four design methods and the baseline strategy.

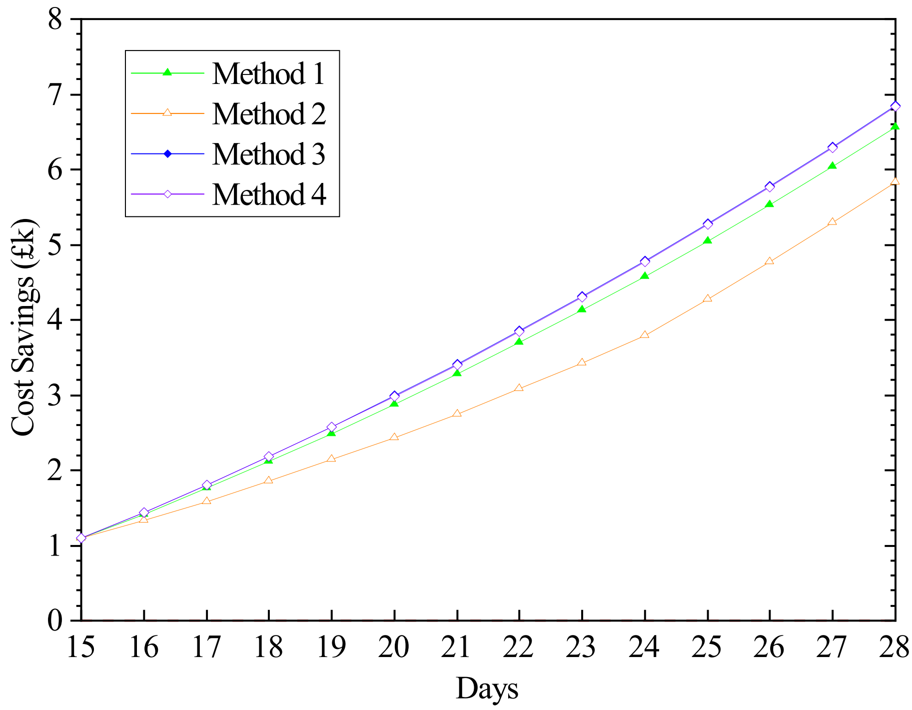


Figure 6.22: Cost Savings of Large Network (Four Design Methods)

To examine the payback period required for achieving the target adoption rate, the modelling calculation is expanded to a longer period without further impose advertisement after Day 28. The daily balance curves for each design method are shown in Fig. 6.23, from which it can be seen that all the methods can reach the stage of having cost benefit over a longer period around 60 days, among them Method 1 has the shortest payback period which is calculated to be Day 64.

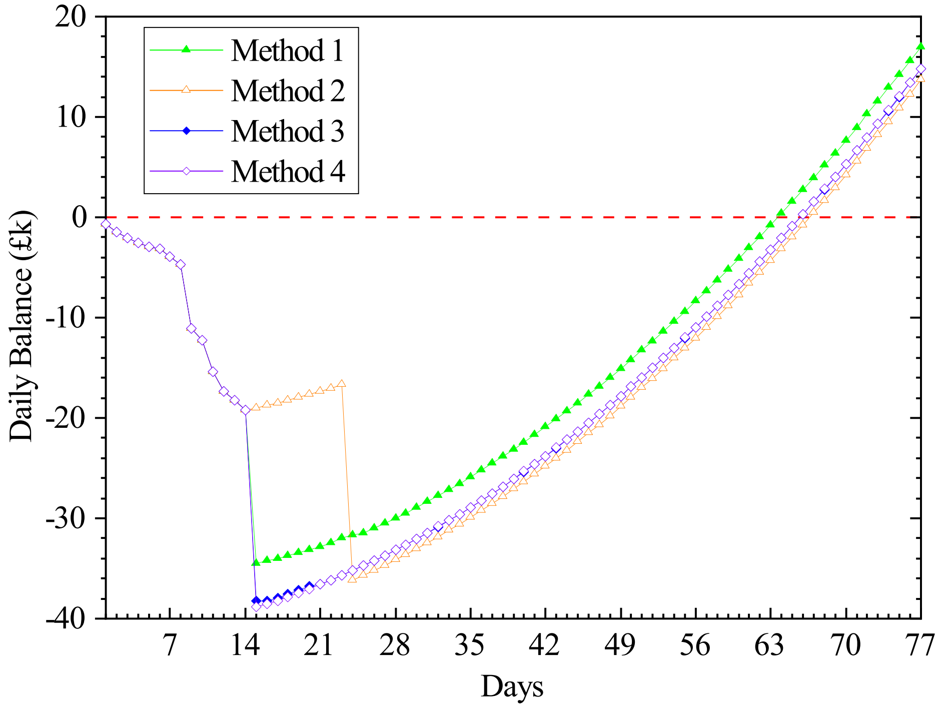


Figure 6.23: Daily Balance of Large Network (Four Design Methods)

6.5 Summary

In this work, a mathematical model has been developed to investigate the impacts of advertisement control to energy savings for population networks. Following the epidemics theory, the small-world network theory, and the BFT, the model is established to describe the dynamics of EEP adoption rate and energy savings subject to advertisement control. The unknown model parameters are estimated by LS identification using survey data. The developed model is applied to a small population network and a large population network for optimal advertisement control design. Sensitivity analysis has been applied to the two network models to examine the possible impact of model uncertainties.

Four optimization design methods are proposed to achieve the target EEP adoption rate under the given budget and the time period. Results from the designed advertisement control strategies are compared with each other and also with the baseline programme that has no advertisement control. The period used for case studies is chosen to be 28 days in total, among which data collected from the first 14 days are used for modelling and the optimization designs are conducted over the last 14 days.

The first design method targets at the lowest advertisement cost. The optimal solutions obtained for both networks demonstrate that, among the four design methods, Method 1 achieves the smallest advertisement investment. See the total impressions over the last 14 days as listed in Table 6.2 for the small network and Table 6.8 for the large network.

With the shortest time objective used in Method 2, the design results show that the large network reaches the targeted adoption rate one day ahead of the full period, and the small network reaches the target four days ahead. Among the four design methods, the results from this method have the shortest time required to achieve the target adoption rates. See Fig. 6.8 for the small network and Fig. 6.17 for the large network.

In the third method with the objective to maximize the total energy savings, the small network has achieved an increase of 32.21% in total energy savings compared to the baseline plan, while for the large population network, it has an increase of 18.37% in energy savings compared to the baseline strategy. The optimized solutions in Method 3 have the largest EES for both small and large networks compared to other methods (see Fig. 6.11 for small network and Fig. 6.20 for large network).

In Method 4 that aims at achieving the largest number of installations, the optimal solutions compared with the baseline strategy lead to an increase of adoption rate of 15.48% and 0.0095% for the small and large population networks, respectively. Compared with the other three design methods, Method 4 obtains the largest adoption rate at the end of the program (see Fig. 6.8 and Fig. 6.17 for small and large networks). The increased levels compared to Methods 1, 2 and 3 are 2.70%, 0.50%, 2.24% for the small network, and $4.92e-5$, $6.71e-6$, $7.22e-8$ for the large network, respectively.

The optimization design studies show that the proposed model can be applied to large-scale mass roll-out programme to forecast the energy savings and predict the impact of advertisement to energy savings. Sensitivity analysis results demonstrate the robustness of the developed model in predicting network adoption rate.

The proposed methodology can be improved from further investigation on modelling and design. In an energy saving scheme applied to a population network, each participant's income, age, education level, etc. can be considered in the modelling as they are related to users' energy consumption behaviours. In practice, a mass roll-out programme usually runs for several years, during which factors such as energy product price variations and holiday sales may significantly affect the adoption rate. These non-routine changes are not included in the current short-period design, but can be involved in a design aiming for energy savings over a longer period of time.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This research contributes to a deeper understanding of behavioral and social factors that influence household energy savings and development of quantitative measures. Based on the objectives set, this PhD work can be summarized into three major contributions.

1. In the first study, the concept of EES is employed to examine the impact of interpersonal relationships on energy consumption within a social network setting.

Introducing node weighting enhances the diversification of connections between individuals, thereby increasing the reliability of quantifying individual influences. Additionally, incorporating a directed network structure provides a more realistic representation of social connections. This study employs complex network modeling to capture the nuances of information diffusion, which includes connection weightings and their evolution over time within a network, thereby creating a dynamic information diffusion system. The primary objective of the model developed for this study is to investigate the impact of household behaviors on energy consumption. It quantifies interactions among individuals and employs EES as a comparable indicator to evaluate the influence of interpersonal dynamics on energy consumption. Using this model, the study also delves into quantifying the EEP adoption information to derive the EES achieved through information diffusion. Sensitivity and cost analyses are included to provide a comprehensive understanding of the EES model's implications and practicalities.

However, it's important to note that the methodologies employed in this study require thorough data exploration, specifically delving into the intricate details of social network connections. Consequently, the applicability of this model is confined to specific social networks and is more suitable for small networks. The model proposed in this study involves numerous network-related parameters that need to be determined, introducing computational challenges when employing the least squares optimization method. The current approach involves selecting a sufficiently large population within a generic algorithm to address the least squares optimization challenge. For larger networks, more advanced optimization algorithms may be necessary.

By analyzing interactions within a social network, exponential functions are utilized to model the decay of information within the network. Leveraging probability theory, the resulting mathematical framework can be used for calculation of EES subject to single information source and multiple source nodes. Furthermore, this model assists in identifying individuals who exert the most substantial impact on the broader social network. To validate the model's effectiveness, it underwent testing within a social network comprising 40 participants, with network data collected through designed survey questionnaires.

Findings from a single-source social network demonstrate that an individual's influence is not solely determined by their connectivity but is also significantly affected by node connection weights. Based on the small network case study system, optimal selection of an information source can result in approximately 25% energy savings in the initial two months, whereas a randomly chosen node yields around 5.64% energy savings. As a result, the model's ability to select optimal nodes can lead to significant additional energy savings within a short time frame. In scenarios where multiple information sources exist within the network, the influence of these sources depends on network connections, connection strength and numerous combination of multiple nodes, rendering the optimization problem more complex.

Additionally, it's worth noting that the model's outcomes are sensitive to changes in network structure, necessitating periodic updates to maintain accuracy of model prediction.

2. In the second study, a mathematical model has been developed which is capable of simulating the adoption rates of energy-efficient products within both small and large network populations, when individual user factors, social network influence and advertisement influence are considered.

This model encompasses four distinct facets of network influence, enabling a comparative analysis of the effects of individual acceptance, social influence, network trend influence, and advertising influence on adoption rates. Through case studies conducted on varying network sizes, the model is utilized to compute the utility coefficient based on survey data. Among the four individual factors pertaining to personal acceptance influence, family and employment status emerged as the most influential determinants impacting the adoption of energy-efficient products.

For both small and large network populations, the primary driver of network influence is social trend influence (overall network adoption rate). This implies that the adoption behavior of households is primarily influenced by the collective adoption rate within the entire social network. In networks with small populations, personal acceptance contributes to 0.2388 of the four utility weightings. However, within larger network populations, personal acceptance demonstrates a reduced impact among households in networks closely resembling aggregated networks. Nonetheless, with an increasing rewire probability, personal acceptance regained prominence, accounting for approximately 0.2386. In practical scenarios, real-world social networks tend to mirror W-S small worlds, with a rewire probability of around 0.1. Consequently, the influence of personal acceptance within larger network populations appears less important compared to its influence within smaller network populations.

3. The third study investigates the impact of optimal advertisement control to adoption of EEP and consequently energy savings and cost benefits.

By integrating methods of epidemic theory, small-world network theory, and Bayesian forecasting theory, a robust mathematical model has been developed for the first time. This model stands out in its ability to quantify advertisement influence to EEP adoption rate, energy savings and cost reduction. The expedition commences with a smaller network to construct a technical framework essential for modeling and optimization design, laying down a strong foundation to be expanded to larger scale networks with reliability and numerical efficiency.

Four divergent optimization design methods are proposed with an aim to reach a target adoption rate within the given budget and time period. A comparative analysis is carried out to understand the benefits of each design as compared to each other and the baseline strategy. The results demonstrate the model's capability and effectiveness in batch roll-out endeavors, opening up avenues to not only forecast energy savings but also to gauge the impact of advertising on promoting energy products and achieving energy savings with a social network.

As the study broadened its horizon to encompass larger population networks, a new insight is obtained – the potency of advertising control magnified, overshadowing social influence which seemed to hold a stronger sway in smaller population networks. The comparative analysis underscores that for larger population networks, advertising shows a stronger influence, as compared to social influence. This amplified impact of advertising holds profound implications for policy design and strategies aimed at fostering energy conservation in larger population networks, underscoring the pivotal role of well-strategized advertising campaigns in nudging the masses towards more sustainable energy consumption behaviors. Hence, as we venture into a future where the dynamics of large population networks continue to evolve, harnessing the power of advertising control emerges as a paramount strategy to drive the narrative of energy conservation forward, offering a lens to look beyond the superficial and delve deeper into the mechanics of influence in large scale population networks.

The work in Chapter 4-6 dive deeper into practical applications and methodologies for energy saving, contrasting with works in Chapter 2's focus on theoretical frameworks and the role of human behavior and social networks. These later chapters present specific case studies, data analysis techniques, and technological interventions designed to improve energy efficiency. They move from the conceptual to the empirical, showcasing how the theories discussed in this thesis can be implemented through engineering practices and models. This progression from theory to application underscores the thesis's holistic approach to addressing energy efficiency, combining insights from social sciences with engineering innovations.

7.2 Future Work

The research presented in this thesis lays the foundation for several potential avenues of extension and improvement. Here are some ideas for advancing this work:

- **Enhanced Data Collection:** To further enhance the model's robustness and applicability, consider expanding the scope of data collection across both micro and macro scales of networks. This expansion should encompass a more diverse participant base, including individuals from various demographic backgrounds, socio-economic statuses, educational levels, and other relevant personal attributes. Increasing the sample size can also help reduce biases and provide a more comprehensive understanding of network behaviors.

- **Scenario Complexity:** Introducing more complex scenarios for testing the model's robustness can provide valuable insights into its limitations and areas for improvement. Explore scenarios characterized by higher levels of uncertainty and multi-dimensional interactions.

- **Behavioral Simulation Techniques:** While the current research employs a probabilistic approach to simulate human behavior, consider exploring alternative methodologies. Incorporating psychological theories and principles could offer a more nuanced understanding of human behavior within networks. Additionally, investigating alternative mathematical or computational frameworks, such as agent-based modeling, behavioral economics models, or machine learning algorithms, may provide different perspectives and potentially improve predictive accuracy.

- **Technological Advancements:** Leverage emerging technologies like advanced analytics, big data, and ML (machine learning) to enhance data analysis and modeling capabilities. A cross-disciplinary approach, integrating insights from psychology, sociology, economics, and other relevant fields, can enrich the model's theoretical foundation and practical applicability. This interdisciplinary approach can foster a more holistic understanding of the intricate interactions within networks.

7.2.1 Potential Use of ML

Replacing traditional model parameter estimation with machine learning for influence utility analysis is not only possible but offers several advantages in terms of flexibility, scalability, and adaptability. However, it's important to be mindful of the challenges, particularly regarding interpretability and data dependency. Successful application requires careful model selection, rigorous validation, and ongoing monitoring to ensure the model remains accurate over time.

Influence utility, in Chapter6, refers to the likelihood or utility of an individual adopting a new behavior, product, or technology based on various influencing factors. The proposal to use ML in this scenario involves taking these factors as inputs and predicting the updated adoption state as the output.

- Model Training: During the training phase, the ML model is exposed to a dataset consisting of various instances where the influencing factors are known, and the corresponding adoption state is also known. The model learns the relationship between these factors and the adoption outcome.

- Model Testing and Use: After training, the model is tested with a separate set of data to evaluate its performance, i.e., its ability to predict the adoption state based on the influencing factors. Once validated, the model can be used in real-world scenarios to predict the adoption state given new sets of influencing factors.

Appendix A

Large Population Network Campaign Data (Population 1 Million)

Date	Impression	Visitor	Deal amount	Visitor/Impression	Deal/Impression	Deal/Visitor	Deal/Population	Visitor/Population
03/12/14	7,246,843	8522	0	0.117596%	0	0.000000%	0	0.008522
04/12/14	7,316,070	5,775	6	0.078936%	8.20112E-07	0.103896%	0.000006	0.005775
05/12/14	6,506,708	5,565	9	0.085527%	1.38319E-06	0.161725%	0.000009	0.005565
06/12/14	4,250,657	5,109	4	0.120193%	9.41031E-07	0.078293%	0.000004	0.005109
07/12/14	4,156,518	4,291	7	0.103235%	1.6841E-06	0.163132%	0.000007	0.004291
08/12/14	2,767,515	4,337	5	0.156711%	1.80667E-06	0.115287%	0.000005	0.004337
09/12/14	7,798,623	11,266	16	0.144461%	2.05164E-06	0.142020%	0.000016	0.011266
10/12/14	7,358,586	6,825	10	0.092749%	1.35896E-06	0.146520%	0.00001	0.006825
11/12/14	64,900,395	112,905	105	0.173967%	1.61786E-06	0.092999%	0.000105	0.112905
12/12/14	12,851,081	7,172	71	0.055809%	5.52483E-06	0.989961%	0.000071	0.007172
13/12/14	32,333,011	13,269	64	0.041039%	1.9794E-06	0.482327%	0.000064	0.013269
14/12/14	21,386,124	12,893	65	0.060287%	3.03935E-06	0.504150%	0.000065	0.012893
15/12/14	10,676,755	8,681	65	0.081307%	6.08799E-06	0.748762%	0.000065	0.008681
16/12/14	11,258,100	7,176	59	0.063741%	5.24067E-06	0.822185%	0.000059	0.007176
17/12/14	8,626,184	13,052	60	0.151307%	6.95557E-06	0.459700%	0.00006	0.013052
18/12/14	11,157,634	14,460	87	0.129597%	7.79735E-06	0.601660%	0.000087	0.01446
19/12/14	10,989,131	18,621	52	0.169449%	4.73195E-06	0.279255%	0.000052	0.018621
20/12/14	10,986,393	15,083	28	0.137288%	2.54861E-06	0.185639%	0.000028	0.015083
21/12/14	12,283,494	15,250	40	0.124150%	3.2564E-06	0.262295%	0.00004	0.01525
22/12/14	9,651,261	15,296	67	0.158487%	6.9421E-06	0.438023%	0.000067	0.015296
23/12/14	10,605,578	15,872	60	0.149657%	5.6574E-06	0.378024%	0.00006	0.015872
24/12/14	8,586,888	15,302	76	0.178202%	8.8507E-06	0.496667%	0.000076	0.015302
25/12/14	9,531,478	21,465	71	0.225201%	7.449E-06	0.330771%	0.000071	0.021465
26/12/14	10,507,586	19,670	60	0.187198%	5.71016E-06	0.305033%	0.00006	0.01967
27/12/14	12,026,972	14,172	25	0.117835%	2.07866E-06	0.176404%	0.000025	0.014172
28/12/14	9,030,185	18,872	55	0.208988%	6.09068E-06	0.291437%	0.000055	0.018872
29/12/14	10,303,387	15,764	34	0.152998%	3.29989E-06	0.215681%	0.000034	0.015764
30/12/14	10,314,633	15,367	63	0.148983%	6.10783E-06	0.409969%	0.000063	0.015367
31/12/14	5,465,951	11,415	55	0.208838%	1.00623E-05	0.481822%	0.000055	0.011415

Appendix B

Small Population Survey Questionnaire

Questionnaire

The researcher and study

I am Feng Du; my PhD projects need to investigate the interaction of people within a community and how the connections can affect energy saving within that community. The purpose of this survey is to test the theory proposed on how connections among people can affect the energy saving information transfer within the community. This will help determine some parameters and data needed to complete my research. Thank you for your participation.

How to answer the questions in this questionnaire

Please complete the questionnaire by answering all the questions below. Tick where appropriate. The information provided will be used for research purpose only and no names or address will be published. Thank you for your time and help

Part A

Name	
Email Address	

From those people listed, kindly tick one number from **1** to **5** on **the level of closeness** you are to the person. The level of closeness basically depends on the frequency of chatting.

- Leave it blank (e.g.: you never meet them)
- 1 means not close at all (e.g.: you see them around)
- 2 means greet when you see (e.g.: you just greet hello when you see them)
- 3 means you talk (e.g.: you enquire how the person is doing)
- 4 means you are close (e.g.: you see each other or chat online at least once a month)
- 5 means you are very close (e.g.: you can call this person your friend)
-

		Relationship (recently)					**Relationship (half year ago)**					**Relationship (one year ago)**				
Surname	**First Name**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Bao	Jie															
Chen	Guandou															
Chen	Minjiang															
Chen	Dong															
Du	Feng															
Fan	Fulin															
Gao	Gao															
Hong	Qiteng															
Huang	Guangming															
Hu	Yihua															
Jing	Yiran															
Kang	Yuchang															
Li	Sirui															
Li	Ruiqi															
Li	Rui															
Li	Peng															
Lian	Yiqing															
Ni	Xinwen															
Sha	Jin															
Shao	Zhonglei															
Sheng	Bojie															
Su	Xiaoke															
Sun	Yanyi															
Tang	Xiaoqing															
Wang	Yue															
Wang	Su															

			Relationship (recently)						Relationship (half year ago)						Relationship (one year ago)				
Surname	**First Name**		**1**	**2**	**3**	**4**	**5**		**1**	**2**	**3**	**4**	**5**		**1**	**2**	**3**	**4**	**5**
Wang	Xiaoyu																		
Xu	Han																		
Yu	Mengran																		
Zeng	Rong																		
Zhang	Zhongshu																		
Zhang	Jixing																		
Zhang	Wendian																		
Zhao	Yiyi																		
Zhao	Jiawei																		
Zhao	Yu																		
Zhong	Yanni																		
Zhou	Linghe																		
Zhou	Yingjia																		
Zhou	Ding																		

Part B

Let's take LED and CFL light bulbs as examples of energy saving products, then compare to traditional light bulbs. Incandescent light bulbs are still quite popular for residential buildings. The listed light bulbs are in same shape, same colour temperature, and 360 °beam angle.

Product Type	**Price** (£)	**Power** (W)	**Lumen** (Amount of visible light)	**Life Time** (Hours)	**Annual Electricity Bill** (3 hours usage per day)	**Payback Period** (Months)
LED	4.97	4.6	470	15000	£0.81	10 Months
CFL	3.36	9	500	10000	£1.58	7.5 Months
Incandescent	1.13	40	630	2000	£7.01	-

Question 1

From the options listed below, kindly tick one number from **1** to **5** on the level of possibility that you will purchase the energy saving product when you receive above information in the form of advertisement under the condition below.

- 1 means definitely will not buy
- 2 means probably will not buy
- 3 means might/might not buy
- 4 means probably will buy
- 5 means definitely will buy

If you receive the above LED/CFL efficient lighting information from internet (e.g. Facebook, YouTube, Twitter, WeChat, emails, etc.) with the following frequency:	**Possibility**				
	1	**2**	**3**	**4**	**5**
Just a **once** off advertisement					
3/week, and continued for **1 week** only					
1/day, and continued for **1 week** only					
1/week , and continued for **2 weeks** only					
3/week, and continued for **2 weeks** only					
1/day, and continued for **2 weeks** only					
1/week, and continued for **1 month** only					
3/week, and continued for **1 month** only					
1/day, and continued for **1 month** only					

Please tick the answer below. In next questions, multiple options are applicable.

Question 2

If one of the people listed in Part A share the above energy saving information to you, what will be your response?

Option	**Answer**
All my lamps are already LED, and no actions need to taken	
Change the lamps to LED in **next few days**	
Change the lamps to LED in **1 month**	

Change the lamps to LED in **3 months**	
Change the lamps to LED in **6 months**	
I'll accept the information only if I'm close to him/her (information provider)	
Change the lamps if more people tell me about it (write number in **Answer** if yes)	
Share the information to more people	
Share the information to more people after I change the lamps	
I live in an accommodation that do not need to pay electricity cost	
No response	

Part C

Suppose that you own a house which can install Solar Photovoltaic (PV) panels. And the detail of payback period of PV in UK is shown below when taking a 16 X 250W panels for example.

PV			
Installation Fee	**Annual Feed-In & Savings**	**Pay Back Time**	**Profit Over 20 Years**
£6,463	£755	7.5 years	£15,532

Question 3

If one of the people listed in Part A share the above energy saving information to you, what will be your response?

Option	**Answer**
Install it in **1 month**	
Install it in **6 months**	
Install it in **1 year**	
Install it in **2 years**	
I'll accept the information only if I'm close to him/her (information provider)	
I'll install it if more people tell me about it (write number in **Answer** if yes)	
Share the information to more people	
Share the information to more people after I install it	
No response	

Thank you for your participation!

Bibliography

Akbari, H., Pomerantz, M. and Taha, H. (2001), 'Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas', *Sol. Energy* **70**(3), 295–310, doi:10.1016/S0038-092X(00)00089-X.

Albert, R. and Barabási, A.-L. (2002), 'Statistical mechanics of complex networks', *Rev. Mod. Phys.* **74**, 47–97, doi:10.1103/RevModPhys.74.47.

Allcott, H. (2011), 'Social norms and energy conservation', *J. Public Econ.* **95**(9-10), 1082–1095, doi:10.1016/j.jpubeco.2011.03.003.

Andersen, R. V., Toftum, J., Andersen, K. K. and Olesen, B. W. (2009), 'Survey of occupant behaviour and control of indoor environment in danish dwellings', *Energy Build.* **41**(1), 11–16, doi:10.1016/j.enbuild.2008.07.004.

Augenbroe, G. (2004), Trends in building simulation, *in* 'Advanced building simulation', Routledge, pp. 18–38, doi:10.4324/9780203073674-7.

Backstrom, L., Boldi, P., Rosa, M., Ugander, J. and Vigna, S. (2012), Four degrees of separation, *in* 'Proceedings of the 4th Annual ACM Web Science Conference', pp. 33–42, doi:10.1145/2380718.2380723.

Bale, C. S., McCullen, N. J., Foxon, T. J., Rucklidge, A. M. and Gale, W. F. (2014), 'Modeling diffusion of energy innovations on a heterogeneous social network and approaches to integration of real-world data', *Complexity* **19**(6), 83–94, doi:10.1002/cplx.21523.

Barabási, A.-L. (2009), 'Scale-free networks: a decade and beyond', *Science* **325**(5939), 412–413, doi:10.1126/science.1173299.

Barabási, A.-L., Albert, R. and Jeong, H. (2000), 'Scale-free characteristics of random networks: the topology of the world-wide web', *Phys. A: Stat. Mech. Appl.* **281**(1-4), 69–77, doi:10.1016/S0378-4371(00)00018-2.

Barabási, A.-L. and Albert, R. (1999), 'Emergence of scaling in random networks', *Science* **286**(5439), 509–512, doi:10.1126/science.286.5439.509.

Bavaresco, M. V., D'Oca, S., Ghisi, E. and Lamberts, R. (2020), 'Methods used in social sciences that suit energy research: A literature review on qualitative methods to assess the human dimension of energy use in buildings', *Energy Build.* **209**, 109702, doi:10.1016/j.enbuild.2019.109702.

Belaïd, F. and Massié, C. (2023), 'Driving forward a low-carbon built environment: The impact of energy context and environmental concerns on building renovation', *Energy Economics* **124**, 106865, doi:10.1016/j.eneco.2023.106865.

Bell, R. and Dean, P. (1972), 'The structure of vitreous silica: Validity of the random network theory', *Philos. Mag.* **25**(6), 1381–1398, doi:10.1080/14786437208223861.

Berta, M., Christandl, M., Colbeck, R., Renes, J. M. and Renner, R. (2010), 'The uncertainty principle in the presence of quantum memory', *Nat. Phys.* **6**(9), 659–662, doi:10.1038/nphys1734.

Boyd, D. M. and Ellison, N. B. (2007), 'Social network sites: Definition, history, and scholarship', *J. Comput.-Mediat. Commun.* **13**(1), 210–230, doi:10.1111/j.1083-6101.2007.00393.x.

Brandon, G. and Lewis, A. (1999), 'Reducing household energy consumption: A qualitative and quantitative field study', *J. Environ. Psychol.* **19**(1), 75–85, doi:10.1006/jevp.1998.0105.

Callaway, D. S., Newman, M. E., Strogatz, S. H. and Watts, D. J. (2000), 'Network robustness and fragility: Percolation on random graphs', *Phys. Rev. Lett.* **85**(25), 5468, doi:10.1103/PhysRevLett.85.5468.

Carpineto, C. and Romano, G. (1996), 'A lattice conceptual clustering system and its application to browsing retrieval', *Mach. Learn.* **24**(2), 95–122, doi:10.1023/A:1018050230279.

Centola, D. (2010), 'The spread of behavior in an online social network experiment', *Science* **329**(5996), 1194–1197, doi:10.1126/science.1185231.

Chakravarty, P. and Gupta, A. (2013), Impact of energy disaggregation on consumer behavior, UC Berkeley: Behavior, Energy and Climate Change Conference. URL: `https://escholarship.org/content/qt62d3456p/qt62d3456p.pdf`. Accessed: 05-09-2022.

Chen, G., Wang, X. and Li, X. (2014), *Fundamentals of complex networks: models, structures and dynamics*, John Wiley & Sons, doi:10.1002/9781118718124.fmatter.

Chen, S. and Chen, B. (2015), 'Urban energy consumption: different insights from energy flow analysis, input–output analysis and ecological network analysis', *Appl. Energy* **138**, 99–107, doi:10.1016/j.apenergy.2014.10.055.

Chicco, G. (2012), 'Overview and performance assessment of the clustering methods for electrical load pattern grouping', *Energy* **42**(1), 68–80, doi:10.1016/j.energy.2011.12.031.

Chicco, G., Napoli, R. and Piglione, F. (2006), 'Comparisons among clustering techniques for electricity customer classification', *IEEE Trans. Power Syst.* **21**(2), 933–940, doi:10.1109/TPWRS.2006.873122.

Cochran, W. G. (1977), *Sampling techniques*, John Wiley & Sons. URL: `https://www.academia.edu/29684662/Cochran_1977_Sampling_Techniques_Third_Edition`. Accessed: 05-09-2022.

Corsi, S. and Sabelli, C. (2004), General blackout in italy sunday september 28, 2003, h. 03: 28: 00, *in* 'IEEE Power Engineering Society General Meeting, 2004.', IEEE, pp. 1691–1702, doi:10.1109/PES.2004.1373162.

Csardi, G., Nepusz, T. et al. (2006), 'The igraph software package for complex network research', *Int. J. Complex Syst.* **1695**(5), 1–9. URL: https://www.researchgate.net/profile/Gabor-Csardi/publication/221995787_The_Igraph_Software_Package_for_Complex_Network_Research/links/0c96051d301a30f265000000/The-Igraph-Software-Package-for-Complex-Network-Research.pdf. Accessed: 05-09-2022.

Darby, S. (2006), 'The effectiveness of feedback on energy consumption: A review for defra of the literature on metering, billing and direct displays', p. 26. URL: https://www.eci.ox.ac.uk/research/energy/downloads/smart-metering-report.pdf. Accessed: 05-09-2022.

Delzendeh, E., Wu, S., Lee, A. and Zhou, Y. (2017), 'The impact of occupants' behaviours on building energy analysis: A research review', *Renew. Sust. Energ. Rev.* **80**, 1061–1071, doi:10.1016/j.rser.2017.05.264.

Department for Levelling Up, H. and Communities (2021), 'Live tables on dwelling stock (including vacants) 2021'. URL: https://www.gov.uk/government/statistical-data-sets/live-tables-on-dwelling-stock-including-vacants. Accessed: 05-09-2022.

Druckman, A. and Jackson, T. (2008), 'Household energy consumption in the uk: A highly geographically and socio-economically disaggregated model', *Energy Policy* **36**(8), 3177–3192, doi:10.1016/j.enpol.2008.03.021.

Du, F. and Yue, H. (2023), 'Impact analysis of individual and network factors to household energy savings', *IFAC-PapersOnLine* **56**(2), 7084–7089, doi:10.1016/j.ifacol.2023.10.573. 22nd IFAC World Congress.

Du, F., Yue, H. and Zhang, J. (2023), 'Influence of advertisement control to residential energy savings in large networks', *Appl. Energy* **333**, 120661, doi:10.1016/j.apenergy.2023.120661.

Du, F., Zhang, J., Li, H., Yan, J., Galloway, S. and Lo, K. L. (2016), 'Modelling the impact of social network on energy savings', *Appl. Energy* **178**, 56–65, doi:10.1016/j.apenergy.2016.06.014.

Du, J. and Pan, W. (2021), 'Evaluating energy saving behavioral interventions through the lens of social practice theory: A case study in hong kong', *Energy and Buildings* **251**, 111353, doi:10.1016/j.enbuild.2021.111353.

D'Oca, S., Hong, T. and Langevin, J. (2018), 'The human dimensions of energy use in buildings: A review', *Renew. Sust. Energ. Rev.* **81**, 731–742, doi:10.1016/j.rser.2017.08.019.

Economidou, M., Ringel, M., Valentova, M., Castellazzi, L., Zancanella, P., Zangheri, P., Serrenho, T., Paci, D. and Bertoldi, P. (2022), 'Strategic energy and climate policy planning: Lessons learned from european energy efficiency policies', *Energy Policy* **171**, 113225, doi:10.1016/j.enpol.2022.11322.

Eikeland, P. O. and Skjærseth, J. B. (2021), 'The politics of low-carbon innovation: Implementing the european union's strategic energy technology plan', *Energy Research & Social Science* **76**, 102043, doi:10.1016/j.erss.2021.102043.

E.I.S. Department for Business (2021), Average unit costs and fixed costs for electricity for uk regions, Technical report, E. I. S. Department for Business. URL: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/875739/table_224.xlsx`. Accessed: 05-09-2022.

Ekpenyong, U. E., Zhang, J. and Xia, X. (2014), 'Mathematical modelling for the social impact to energy efficiency savings', *Energy Build.* **84**, 344–351, doi:10.1016/j.enbuild.2014.08.019.

Ekpenyong, U. E., Zhang, J. and Xia, X. (2015), 'How information propagation in social networks can improve energy savings based on time of use tariff', *Sustain. Cities Soc.* **19**, 26–33, doi:10.1016/j.scs.2015.07.005.

Ellison, N. B., Steinfield, C. and Lampe, C. (2007), 'The benefits of facebook "friends:" social capital and college students' use of online social network sites', *J. Comput.-Mediat. Commun.* **12**(4), 1143–1168, doi:10.1111/j.1083-6101.2007.00367.x.

Enteria, N. and Mizutani, K. (2011), 'The role of the thermally activated desiccant cooling technologies in the issue of energy and environment', *Renew. Sust. Energ. Rev.* **15**(4), 2095–2122, doi:10.1016/j.rser.2011.01.013.

Estrada, E. (2011), *The structure of complex networks: theory and applications*, Oxford University Press, doi:10.1093/acprof:oso/9780199591756.001.0001.

Fabi, V., Andersen, R. V., Corgnati, S. and Olesen, B. W. (2012), 'Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models', *Build. Environ.* **58**, 188–198, doi:10.1016/j.buildenv.2012.07.009.

Fabi, V., Barthelmes, V. M., Schweiker, M. and Corgnati, S. P. (2017), 'Insights into the effects of occupant behaviour lifestyles and building automation on building energy use', *Energy Procedia* **140**, 48–56, doi:10.1016/j.egypro.2017.11.122.

Feng, Y., Chen, S. and Zhang, L. (2013), 'System dynamics modeling for urban energy consumption and co2 emissions: A case study of beijing, china', *Ecol. Model.* **252**, 44–52, doi:10.1016/j.ecolmodel.2012.09.008.

Figueiredo, V., Rodrigues, F., Vale, Z. and Gouveia, J. B. (2005), 'An electric energy consumer characterization framework based on data mining techniques', *IEEE Trans. Power Syst.* **20**(2), 596–602, doi:10.1109/TPWRS.2005.846234.

Fischer, C. (2008), 'Feedback on household electricity consumption: a tool for saving energy?', *Energy Effic.* **1**(1), 79–104, doi:10.1007/s12053-008-9009-7.

Gao, X. and Malkawi, A. (2014), 'A new methodology for building energy performance benchmarking: An approach based on intelligent clustering algorithm', *Energy Build.* **84**, 607–616, doi:10.1016/j.enbuild.2014.08.030.

George, A. et al. (1970), 'The market for lemons: Quality uncertainty and the market mechanism', doi:10.2307/1879431.

Guerra Santin, O. (2013), 'Occupant behaviour in energy efficient dwellings: evidence of a rebound effect', *J. Hous. Built Environ.* **28**(2), 311–327, doi:10.1007/s10901-012-9297-2.

Guimera, R., Sales-Pardo, M. and Amaral, L. A. N. (2004), 'Modularity from fluctuations in random graphs and complex networks', *Phys. Rev. E* **70**(2), 025101, doi:10.1103/PhysRevE.70.025101.

Gul, M. S. and Patidar, S. (2015), 'Understanding the energy consumption and occupancy of a multi-purpose academic building', *Energy Build.* **87**, 155–165, doi:10.1016/j.enbuild.2014.11.027.

Guy, S. (2006), 'Designing urban knowledge: competing perspectives on energy and buildings', *Environ. Plan. C: Politics Space* **24**(5), 645–659, doi:10.1068/c0607j.

Hakansson, A., Höjer, M., Howlett, R. J. and Jain, L. C. (2013), Recent research in sustainability in energy and buildings, *in* 'Smart Innovation, Systems and Technologies', Vol. 22, Springer, pp. 875–886, doi:10.1007/978-3-642-36645-1.

Harrison, P. J. and Stevens, C. F. (1976), 'Bayesian forecasting', *J. R. Stat. Soc., B: Stat. Methodol.* **38**(3), 205–228, doi:10.1111/j.2517-6161.1976.tb01586.x.

He, R., Jin, J., Qiu, X., Zhang, C. and Yan, J. (2023), 'Rural residents' climate change perceptions, personal experiences, and purchase intention–behavior gap in energy-saving refrigeration appliances in southwest china', *Environmental Impact Assessment Review* **98**, 106967, doi:10.1016/j.eiar.2022.106967.

Hertin, J., Berkhout, F., Gann, D. and Barlow, J. (2003), 'Climate change and the uk house building sector: perceptions, impacts and adaptive capacity', *Build. Res. Inf.* **31**(3-4), 278–290, doi:10.1080/0961321032000097683.

Hopfe, C. J. and Hensen, J. L. (2011), 'Uncertainty analysis in building performance simulation for design support', *Energy Build.* **43**(10), 2798–2805, doi:10.1016/j.enbuild.2011.06.034.

İpek, Ö. and İpek, E. (2022), 'Determinants of energy demand for residential space heating in turkey', *Renewable Energy* **194**, 1026–1033, doi:10.1016/j.renene.2022.05.158.

Irawan, C. A., Jones, D., Hofman, P. S. and Zhang, L. (2023), 'Integrated strategic energy mix and energy generation planning with multiple sustainability criteria and hierarchical stakeholders', *European Journal of Operational Research* **308**(2), 864–883, doi:10.1016/j.ejor.2022.11.044.

Jain, R. K., Gulbinas, R., Taylor, J. E. and Culligan, P. J. (2013), 'Can social influence drive energy savings? detecting the impact of social influence on the energy consumption behavior of networked users exposed to normative eco-feedback', *Energy Build.* **66**, 119–127, doi:10.1016/j.enbuild.2013.06.029.

Jiménez, A., Tiampo, K. and Posadas, A. (2008), 'Small world in a seismic network: the california case', *Nonlinear Processes in Geophysics* **15**(3), 389–395, doi:10.5194/npg-15-389-2008.

Keirstead, J. (2006), 'Evaluating the applicability of integrated domestic energy consumption frameworks in the uk', *Energy policy* **34**(17), 3065–3077, doi:10.1016/j.enpol.2005.06.004.

Kelly, S., Shipworth, M., Shipworth, D., Gentry, M., Wright, A., Pollitt, M., Crawford-Brown, D. and Lomas, K. (2013), 'Predicting the diversity of internal temperatures from the english residential sector using panel methods', *Appl. Energy* **102**, 601–621, doi:10.1016/j.apenergy.2012.08.015.

Klein, L., Kwak, J.-y., Kavulya, G., Jazizadeh, F., Becerik-Gerber, B., Varakantham, P. and Tambe, M. (2012), 'Coordinating occupant behavior for building energy and comfort management using multi-agent systems', *Autom. Constr.* **22**, 525–536, doi:10.1016/j.autcon.2011.11.012.

Kleinfeld, J. S. (2002), 'The small world problem', *Society* **39**(2), 61–66, doi:10.1007/BF02717530.

Knuepfer, K., Rogalski, N., Knuepfer, A., Esteban, M. and Shibayama, T. (2022), 'A reliable energy system for japan with merit order dispatch, high variable renewable share and no nuclear power', *Applied Energy* **328**, 119840, doi:10.1016/j.apenergy.2022.119840.

Kossinets, G. and Watts, D. J. (2006), 'Empirical analysis of an evolving social network', *Science* **311**(5757), 88–90, doi:10.1126/science.1116869.

Krzysztofowicz, R. (1999), 'Bayesian theory of probabilistic forecasting via deterministic hydrologic model', *Water Resour. Res.* **35**(9), 2739–2750, doi:10.1029/1999WR900099.

Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P. et al. (2009), 'Trends in the sources and sinks of carbon dioxide', *Nat. Geosci.* **2**(12), 831–836, doi:10.1038/ngeo689.

Li, Q., Hu, H., Ma, L., Wang, Z., Arıcı, M., Li, D., Luo, D., Jia, J., Jiang, W. and Qi, H. (2022), 'Evaluation of energy-saving retrofits for sunspace of rural residential buildings based on orthogonal experiment and entropy weight method', *Energy for Sustainable Development* **70**, 569–580, doi:10.1016/j.esd.2022.09.007.

Liang, J., Chin, K.-S., Dang, C. and Yam, R. C. (2002), 'A new method for measuring uncertainty and fuzziness in rough set theory', *Int. J. General Syst.* **31**(4), 331–342, doi:10.1080/0308107021000013635.

Lusher, D., Koskinen, J. and Robins, G. (2013), *Exponential random graph models for social networks: Theory, methods, and applications*, Cambridge University Press, doi:10.1017/CBO9780511894701.

Lutzenhiser, L. (1993), 'Social and behavioral aspects of energy use', *Annu. Rev. Environ. Resour.* **18**(1), 247–289, doi:10.1146/annurev.eg.18.110193.001335.

Ma, G., Ma, J., Li, H., Wang, Y., Wang, Z. and Zhang, B. (2022), 'Customer behavior in purchasing energy-saving products: Big data analytics from online reviews of e-commerce', *Energy Policy* **165**, 112960, doi:10.1016/j.enpol.2022.112960.

Ma, Y. and Liu, C. (2023), 'Emotional or rational choice: The influence of individual personality on energy-saving behavior', *Energy Economics* p. 106768, doi:10.1016/j.eneco.2023.106768.

Mabogunje, A. L. (1970), 'Systems approach to a theory of rural-urban migration', *Geogr. Anal.* **2**(1), 1–18, doi:10.1111/j.1538-4632.1970.tb00140.x.

Mankoff, J., Matthews, D., Fussell, S. R. and Johnson, M. (2007), Leveraging social networks to motivate individuals to reduce their ecological footprints, *in* '2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)', IEEE, pp. 87–87, doi:10.1109/HICSS.2007.325.

Mansur, V., Carreira, P. and Arsenio, A. (2014), A learning approach for energy efficiency optimization by occupancy detection, *in* 'International Internet of Things Summit', Springer, pp. 9–15, doi:10.1007/978-3-319-19656-5_2.

Masoso, O. T. and Grobler, L. J. (2010), 'The dark side of occupants' behaviour on building energy use', *Energy Build.* **42**(2), 173–177, doi:10.1016/j.enbuild.2009.08.009.

Mazur-Stommen, S. and Farley, K. (2013), Aceee field guide to utility-run behavior programs, American Council for an Energy-Efficient Economy. URL: `https://c2e2.unepccc.org/wp-content/uploads/sites/3/2016/03/b132.pdf`. Accessed: 05-09-2022.

McCullen, N. J., Rucklidge, A. M., Bale, C. S., Foxon, T. J. and Gale, W. F. (2013), 'Multiparameter models of innovation diffusion on complex networks', *SIAM J. Appl. Dyn.* **12**(1), 515–532, doi:10.1137/120885371.

Meakin, P. (1983), 'Formation of fractal clusters and networks by irreversible diffusion-limited aggregation', *Phys. Rev. Lett.* **51**(13), 1119, doi:10.1103/PhysRevLett.51.1119.

Moezzi, M. and Janda, K. B. (2014), 'From "if only" to "social potential" in schemes to reduce building energy use', *Energy Res. Soc. Sci.* **1**, 30–40, doi:10.1016/j.erss.2014.03.014.

Mohseni, S., Brent, A. C., Kelly, S., Browne, W. N. and Burmester, D. (2021), 'Strategic design optimisation of multi-energy-storage-technology micro-grids considering a two-stage game-theoretic market for demand response aggregation', *Applied Energy* **287**, 116563, doi:10.1016/j.apenergy.2021.116563.

Newman, M. E. (2003), 'The structure and function of complex networks', *SIAM Rev.* **45**(2), 167–256, doi:10.1137/s003614450342480.

Newman, M. E. and Watts, D. J. (1999*a*), 'Renormalization group analysis of the small-world network model', *Phys. Lett. A* **263**(4-6), 341–346, doi:10.1016/S0375-9601(99)00757-4.

Newman, M. E. and Watts, D. J. (1999*b*), 'Scaling and percolation in the small-world network model', *Phys. Rev. E* **60**(6), 7332, doi:10.1103/PhysRevE.60.7332.

Nguyen, A.-T., Reiter, S. and Rigo, P. (2014), 'A review on simulation-based optimization methods applied to building performance analysis', *Appl. Energy* **113**, 1043–1058, doi:10.1016/j.apenergy.2013.08.061.

NOP, G. and Research, K. M. (2012), 'Research report green deal segmentation'. URL: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/49750/Green_Deal_segmentation_-_research_report.pdf`. Accessed: 05-09-2022.

Owens, S. and Driffill, L. (2008), 'How to change attitudes and behaviours in the context of energy', *Energy policy* **36**(12), 4412–4418, doi:10.1016/j.enpol.2008.09.031.

Paauw, J., Roossien, B., Aries, M. and Santin, O. G. (2009), 'Energy pattern generator understanding the effect of user behaviour on energy systems'. URL:

https://www.eceee.org/static/media/uploads/site-2/library/conference_proceedings/EE_and_Behaviour/2009/Panel_2/2.206/paper.pdf. Accessed: 05-09-2022.

Palmer, J. and Cooper, I. (2013), 'United kingdom housing energy fact file'. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/345141/uk_housing_fact_file_2013.pdf. Accessed: 05-09-2022.

Pan, S., Wang, X., Wei, Y., Zhang, X., Gal, C., Ren, G., Yan, D., Shi, Y., Wu, J., Xia, L. et al. (2017), Cluster analysis for occupant-behavior based electricity load patterns in buildings: A case study in shanghai residences, *in* 'Building simulation', Vol. 10, Springer, pp. 889–898, doi:10.1007/s12273-017-0377-9.

Parker, S. (2021), Energy consumption in the uk (ecuk): Final energy consumption tables 2021, Technical report. URL: https://www.gov.uk/government/statistics/energy-consumption-in-the-uk-2021. Accessed: 05-09-2022.

Peng, C., Yan, D., Wu, R., Wang, C., Zhou, X. and Jiang, Y. (2012), Quantitative description and simulation of human behavior in residential buildings, *in* 'Building simulation', Vol. 5, Springer, pp. 85–94, doi:10.1007/s12273-011-0049-0.

Pérez-Lombard, L., Ortiz, J. and Pout, C. (2008), 'A review on buildings energy consumption information', *Energy Build.* **40**(3), 394–398, doi:10.1016/j.enbuild.2007.03.007.

Peschiera, G. and Taylor, J. E. (2012), 'The impact of peer network position on electricity consumption in building occupant networks utilizing energy feedback systems', *Energy Build.* **49**, 584–590, doi:10.1016/j.enbuild.2012.03.011.

Pole, A., West, M. and Harrison, J. (2018), *Applied Bayesian forecasting and time series analysis*, Chapman and Hall/CRC, doi:10.1201/9781315274775.

Poortinga, W., Steg, L. and Vlek, C. (2004), 'Values, environmental concern, and environmental behavior: A study into household energy use', *Environ. Behav.* **36**(1), 70–93, doi:10.1177/0013916503251466.

Qin, H. and Pan, W. (2020), 'Energy use of subtropical high-rise public residential buildings and impacts of energy saving measures', *J. Clean. Prod.* **254**, 120041, doi:10.1016/j.jclepro.2020.120041.

Rees, W. and Wackernagel, M. (2008), Urban ecological footprints: why cities cannot be sustainable—and why they are a key to sustainability, *in* 'Urban ecology', Springer, pp. 537–555, doi:10.1007/978-0-387-73412-5_35.

Rozhkov, A. (2023), 'Harnessing european policies for energy planning in illinois: Overcoming barriers and transitioning to a climate-neutral society', *Sustainable Cities and Society* **98**, 104803, doi:10.1016/j.scs.2023.104803.

Rubinov, M. and Sporns, O. (2010), 'Complex network measures of brain connectivity: uses and interpretations', *Neuroimage* **52**(3), 1059–1069, doi:10.1016/j.neuroimage.2009.10.003.

Ruth, M. and Coelho, D. (2015), Understanding and managing the complexity of urban systems under climate change, *in* 'Integrating climate change actions into local development', Routledge, pp. 317–336, doi:10.4324/9781849771757.

Sartori, I. and Hestnes, A. G. (2007), 'Energy use in the life cycle of conventional and low-energy buildings: A review article', *Energy Build.* **39**(3), 249–257, doi:10.1016/j.enbuild.2006.07.001.

Scott, J. (2017), *Social Network Analysis*, SAGE publications, doi:10.4135/9781529716597.

Shannon, C. E. (1948), 'A mathematical theory of communication', *Bell Syst. Tech. J.* **27**(3), 379–423, doi:10.1002/j.1538-7305.1948.tb01338.x.

Shapiro, C. (1983), 'Premiums for high quality products as returns to reputations', *Q. J. Econ.* **98**(4), 659–679, doi:10.2307/1881782.

Sillak, S., Borch, K. and Sperling, K. (2021), 'Assessing co-creation in strategic planning for urban energy transitions', *Energy Research & Social Science* **74**, 101952, doi:10.1016/j.erss.2021.101952.

Simanaviciene, Z., Volochovic, A., Vilke, R., Palekiene, O. and Simanavicius, A. (2015), 'Research review of energy savings changing people's behavior: a case of foreign country', *Procedia Soc. Behav. Sci.* **191**, 1996–2001, doi:10.1016/j.sbspro.2015.04.315.

Srinivasa, S. and Haenggi, M. (2009), 'Distance distributions in finite uniformly random networks: Theory and applications', *IEEE Trans. Veh. Technol.* **59**(2), 940–949, doi:10.1109/TVT.2009.2035044.

Stankovic, L., Stankovic, V., Liao, J. and Wilson, C. (2016), 'Measuring the energy intensity of domestic activities from smart meter data', *Applied Energy* **183**, 1565–1580, doi:10.1016/j.apenergy.2016.09.087.

Stauffer, D. (1979), 'Scaling theory of percolation clusters', *Phys. Rep.* **54**(1), 1–74, doi:10.1016/0370-1573(79)90060-7.

Stutzman, F. (2006), 'An evaluation of identity-sharing behavior in social network communities', *J. Int. Digit. Media Arts Assoc.* **3**(1), 10–18. URL: `http://fredstutzman.com.s3.amazonaws.com/papers/IDMA2006_Stutzman.pdf`. Accessed: 05-09-2022.

Sun, K. and Hong, T. (2017), 'A simulation approach to estimate energy savings potential of occupant behavior measures', *Energy Build.* **136**, 43–62, doi:10.1016/j.enbuild.2016.12.010.

Swan, L. G. and Ugursal, V. I. (2009), 'Modeling of end-use energy consumption in the residential sector: A review of modeling techniques', *Renew. Sust. Energ. Rev.* **13**(8), 1819–1835, doi:10.1016/j.rser.2008.09.033.

Taneja, S. and Mandys, F. (2022), 'Drivers of uk household energy expenditure: Promoting efficiency and curbing emissions', *Energy Policy* **167**, 113042, doi:10.1016/j.enpol.2022.113042.

Travers, J. and Milgram, S. (1977), An experimental study of the small world problem, *in* 'Social networks', Elsevier, pp. 179–197, doi:10.1016/B978-0-12-442450-0.50018-3.

Ulanowicz, R. E., Goerner, S. J., Lietaer, B. and Gomez, R. (2009), 'Quantifying sustainability: Resilience, efficiency and the return of information theory', *Ecol. Complex.* **6**(1), 27–36, doi:10.1016/j.ecocom.2008.10.005.

Vakiloroaya, V., Samali, B., Fakhar, A. and Pishghadam, K. (2014), 'A review of different strategies for hvac energy saving', *Energy Convers. Manag.* **77**, 738–754, doi:10.1016/j.enconman.2013.10.023.

Valente, T. W. (2010), *Social networks and health: Models, methods, and applications*, Oxford University Press, doi:10.1093/acprof:oso/9780195301014.001.0001.

Van Houwelingen, J. H. and Van Raaij, W. F. (1989), 'The effect of goal-setting and daily electronic feedback on in-home energy use', *J. Consum. Res.* **16**(1), 98–105, doi:10.1086/209197.

Van Raaij, W. F. and Verhallen, T. M. (1983*a*), 'A behavioral model of residential energy use', *J. Econ. Psychol.* **3**(1), 39–63, doi:10.1016/0167-4870(83)90057-0.

Van Raaij, W. F. and Verhallen, T. M. (1983*b*), 'Patterns of residential energy behavior', *J. Econ. Psychol.* **4**(1-2), 85–106, doi:10.1016/0167-4870(83)90047-8.

Verderber, R. and Rubinstein, F. (1984), 'Mutual impacts of lighting controls and daylighting applications', *Energy Build.* **6**(2), 133–140, doi:10.1016/0378-7788(84)90068-9.

Wasserman, S. (1994), *Advances in social network analysis: Research in the social and behavioral sciences*, Sage, doi:10.4135/9781452243528.

Watts, D. J. and Strogatz, S. H. (1998), 'Collective dynamics of 'small-world'networks', *Nature* **393**(6684), 440–442, doi:10.1038/30918.

Weimann, J. (1994), 'Individual behaviour in a free riding experiment', *J. Public Econ.* **54**(2), 185–200, doi:10.1016/0047-2727(94)90059-0.

Weinstein, R., Scott, R. and Jones, C. (1989), 'Measurement of "free-riders" in energy conservative programs', *Eval. Program Plan.* **12**(2), 121–130, doi:10.1016/0149-7189(89)90002-5.

Wellmann, J. F. and Regenauer-Lieb, K. (2012), 'Uncertainties have a meaning: Information entropy as a quality measure for 3-d geological models', *Tectonophysics* **526**, 207–216, doi:10.1016/j.tecto.2011.05.001.

West, M. and Harrison, J. (1997), *Bayesian forecasting and dynamic models*, Springer, doi:10.1007/b98971.

Wilhite, H., Nakagami, H., Masuda, T., Yamaga, Y. and Haneda, H. (1996), 'A cross-cultural analysis of household energy use behaviour in japan and norway', *Energy policy* **24**(9), 795–803, doi:10.1016/0301-4215(96)00061-4.

Williamson, T. J. (2010), 'Predicting building performance: the ethics of computer simulation', *Build. Res. Inf.* **38**(4), 401–410, doi:10.1080/09613218.2010.481204.

Wolfe, A. K., Malone, E. L., Heerwagen, J. H. and Dion, J. P. (2014), Behavioral change and building performance: strategies for significant, persistent, and measurable institutional change, Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Xu, X., Taylor, J. E., Pisello, A. L. and Culligan, P. J. (2012), 'The impact of place-based affiliation networks on energy conservation: An holistic model that integrates the influence of buildings, residents and the neighborhood context', *Energy Build.* **55**, 637–646, doi:10.1016/j.enbuild.2012.09.013.

Yan, D., O'Brien, W., Hong, T., Feng, X., Gunay, H. B., Tahmasebi, F. and Mahdavi, A. (2015), 'Occupant behavior modeling for building performance simulation: Current state and future challenges', *Energy Build.* **107**, 264–278, doi:10.1016/j.enbuild.2015.08.032.

Yao, R. and Steemers, K. (2005), 'A method of formulating energy load profile for domestic buildings in the uk', *Energy Build.* **37**(6), 663–671, doi:10.1016/j.enbuild.2004.09.007.

Ye, X., Xia, X. and Zhang, J. (2014), 'Optimal sampling plan for clean development mechanism lighting projects with lamp population decay', *Appl. Energy* **136**, 1184–1192, doi:10.1016/j.apenergy.2014.07.056.

Yohanis, Y. G., Mondol, J. D., Wright, A. and Norton, B. (2008), 'Real-life energy use in the uk: How occupancy and dwelling characteristics affect domestic electricity use', *Energy Build.* **40**(6), 1053–1059, doi:10.1016/j.enbuild.2007.09.001.

Yu, W., Chen, G. and Lü, J. (2009), 'On pinning synchronization of complex dynamical networks', *Automatica* **45**(2), 429–435, doi:10.1016/j.automatica.2008.07.016.

Yu, Z., Fung, B. C., Haghighat, F., Yoshino, H. and Morofsky, E. (2011), 'A systematic procedure to study the influence of occupant behavior on building energy consumption', *Energy Build.* **43**(6), 1409–1417, doi:10.1016/j.enbuild.2011.02.002.

Zhang, H., Wang, H. and Zhou, X. (2017), Applicability research on passive design of residential buildings in hot summer and cold winter zone in china, *in* 'IOP Conference Series: Earth and Environmental Science', Vol. 61, IOP Publishing, p. 012066, doi:10.1088/1755-1315/61/1/012066.

Zhang, Y., Bai, X., Mills, F. P. and Pezzey, J. C. (2018), 'Rethinking the role of occupant behavior in building energy performance: A review', *Energy Build.* **172**, 279–294, doi:10.1016/j.enbuild.2018.05.017.

Zhao, D., McCoy, A. P., Du, J., Agee, P. and Lu, Y. (2017), 'Interaction effects of building technology and resident behavior on energy consumption in residential buildings', *Energy Build.* **134**, 223–233, doi:10.1016/j.enbuild.2016.10.049.

Zhao, H.-x. and Magoulès, F. (2012), 'A review on the prediction of building energy consumption', *Renew. Sust. Energ. Rev.* **16**(6), 3586–3592, doi:10.1016/j.rser.2012.02.049.

Zhao, J., Lasternas, B., Lam, K. P., Yun, R. and Loftness, V. (2014), 'Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining', *Energy Build.* **82**, 341–355, doi:10.1016/j.enbuild.2014.07.033.