EAI/Springer Innovations in Communication and Computing

Suman Paul
Sara Paiva
Bo Fu *Editors*

Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services

Springer

EAI/Springer Innovations in Communication and Computing

Series Editor

Imrich Chlamtac, European Alliance for Innovation, Ghent, Belgium

Editor's Note

The impact of information technologies is creating a new world yet not fully understood. The extent and speed of economic, life style and social changes already perceived in everyday life is hard to estimate without understanding the technological driving forces behind it. This series presents contributed volumes featuring the latest research and development in the various information engineering technologies that play a key role in this process.

The range of topics, focusing primarily on communications and computing engineering include, but are not limited to, wireless networks; mobile communication; design and learning; gaming; interaction; e-health and pervasive healthcare; energy management; smart grids; internet of things; cognitive radio networks; computation; cloud computing; ubiquitous connectivity, and in mode general smart living, smart cities, Internet of Things and more. The series publishes a combination of expanded papers selected from hosted and sponsored European Alliance for Innovation (EAI) conferences that present cutting edge, global research as well as provide new perspectives on traditional related engineering fields. This content, complemented with open calls for contribution of book titles and individual chapters, together maintain Springer's and EAI's high standards of academic excellence. The audience for the books consists of researchers, industry professionals, advanced level students as well as practitioners in related fields of activity include information and communication specialists, security experts, economists, urban planners, doctors, and in general representatives in all those walks of life affected ad contributing to the information revolution.

Indexing: This series is indexed in Scopus, Ei Compendex, and zbMATH.

About EAI

EAI is a grassroots member organization initiated through cooperation between businesses, public, private and government organizations to address the global challenges of Europe's future competitiveness and link the European Research community with its counterparts around the globe. EAI reaches out to hundreds of thousands of individual subscribers on all continents and collaborates with an institutional member base including Fortune 500 companies, government organizations, and educational institutions, provide a free research and innovation platform.

Through its open free membership model EAI promotes a new research and innovation culture based on collaboration, connectivity and recognition of excellence by community.

More information about this series at http://www.springer.com/series/15427

Suman Paul • Sara Paiva • Bo Fu
Editors

Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services

Springer

EAI
RESEARCH MEETS INNOVATION

Editors
Suman Paul
Department of ECE
Haldia Institute of Technology
Haldia, West Bengal, India

Sara Paiva
Inst Politécnico de Viana do Castelo
Viana do Castelo, Portugal

Bo Fu
Computer Engineering
and Computer Science
California State University, Long Beach
Long Beach, CA, USA

ISSN 2522-8595 ISSN 2522-8609 (electronic)
EAI/Springer Innovations in Communication and Computing
ISBN 978-3-030-77557-5 ISBN 978-3-030-77558-2 (eBook)
https://doi.org/10.1007/978-3-030-77558-2

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book intends to gather valuable contributions to help the reader to understand the latest technological developments in the areas of cutting edge developments of next generation data and knowledge management for the diversified applications related to the convergence of ICT, Healthcare and telecommunication services.

The first chapter significantly introduces and focuses the scope of this book by addressing an interoperable and integrated care service architecture for intellectual disability services in the context of the COVID-19 pandemic. The following two chapters focus on the work related to big data in the context of healthcare and in applications related to intelligent transport systems. The next chapter provides an in-depth study on the applications of intelligent sensors in healthcare services. Finally, the last two chapter address the convergence of ICT, which deal with the applications of robotic arm with ANN and IOT, respectively.

This book attracted contributors from several countries. Therefore, we would like to thank all authors for their contributions and also thank all reviewers for their review work. Last but not least, we express our sincerest thanks to Dr. Imrich Chlamtac, Series Editor, and Ms. Mary E. James, Senior Editor, EAI/Springer. We convey special thanks to Ms. Eliška Vlˇcková, Managing Editor, European Alliance for Innovation (EAI), for all her support and cooperation throughout the process of creating this book.

Haldia, West Bengal, India — Suman Paul
Viana do Castelo, Portugal — Sara Paiva
Long Beach, CA, USA — Bo Fu

Contents

About the Editors

Suman Paul is currently an assistant professor in the Department of Electronics and Communication Engineering at Haldia Institute of Technology, India. His interests lie in scheduling, QoS in communication networks, application of soft computing techniques in networking, and applications of IoT. Professor Paul worked as associate researcher at the Indian Institute of Management (IIM) Calcutta in a project on Zigbee-based wireless sensor network, funded by DIT, Govt. of India. He is a regular reviewer of SCIE/Scopus indexed journals. Professor Paul has produced books on mobile ad hoc network; published research articles in reputed journals by publishers such as Springer and the Korean Society for Internet Information; written book chapters for Springer; and presented papers in IEEE International conferences. He was the joint convener of the International Conference, ICCDC 2019, Haldia, India. He is a member of ACM, Machine Intelligence Research Labs, USA, and qualified Cisco Certified Network Associate.

Sara Paiva is an assistant professor at the Polytechnic Institute of Viana do Castelo, a Ph.D. in informatics engineering from the University of Vigo (2011), and a postdoctoral researcher at the University of Oviedo, under advanced driving assistants. Her main line of research is smart inclusive mobility in smart cities. Sara is currently vice chair of the IEEE Smart Cities Marketing Committee, editor-in-chief of *EAI Endorsed Transactions on Smart Cities*, associate editor in several IF journals such as *MONET* and *WINET*, and editorial board member of the *5.2 IF International Journal of Sustainable Cities and Society*. She is general co-chair of the EAI International Convention on Smart Cities 360° since the 2020 edition and is part of the Organizing Committee of IEEE Smart Cities 2021. She is scientific evaluator of Individual Scientist Project Proposals for Fundamental and Applied Research for the Latvian Council of Science, and is founding fellow of AI Live Talks, in partnership with IEEE. Sara is also editor of books with Springer, CRC Press, Cambridge Scholar, and special issues in several impact factor journals. She has authored and co-authored several scientific publications in journals and conferences and is a frequent reviewer of international journals and international conferences.

Bo Fu is an assistant professor in the Department of Computer Engineering and Computer Science at California State University Long Beach (CSULB), California, USA. Dr. Fu received her B.Sc. (Joint Hons.), M.Sc., and Ph.D. in computer science from Trinity College Dublin, Ireland. Prior to joining CSULB, Dr. Fu was a research scientist at the National Center for Biomedical Ontology, and a postdoctoral fellow at the University of Victoria, Canada.

Dr. Fu's research intersects Web semantics and human data interaction. She is interested in accelerating research and innovation in semantic-oriented intelligent systems and interactive visual analytics that facilitate human decision making. Through semantics and interaction, the overall goal of her research is to help people make sense of complex datasets during the process of turning data to information, to knowledge, and to wisdom in collective intelligence.

Dr. Fu has published her research at premier journals and conferences including the *Semantic Web Journal*, the *Journal of Web Semantics*, the *International Semantic Web Conference*, and the Extended Semantic Web Conference. She has chaired and reviewed numerous scientific panels and committees, such as National Science Foundation programs, Big Data Research, Artificial Intelligence in Medicine, and BMC Bioinformatics.

Chapter 1
Development of an Interoperable-Integrated Care Service Architecture for Intellectual Disability Services: An Irish Case Study

Subhashis Das and Pamela Hussey

1.1 Introduction

There is much evidence to suggest that digital transformations are complex processes that are predisposed to fail. Only about one in four transformations succeeded in the short and long terms, and the success rate has been trending downward [4, 7, 54]. Policy makers and administrators charged with digital transformation at the enterprise level need to consider carefully corner stones on specific design approaches. For example, what judgements they make and what they base their core decision making on. Important decisions need to be influenced by a critical review of emerging evidence, what data analytics requirements are needed, with some focused consideration on how emerging technology can provide support at policy clinical and administration levels in short, medium, and long terms [16, 17, 22, 33]. The world is changing and we are now living in a post COVID-19 society. This is having an impact on how we live and how we work. Telehealth is emerging as key driver to support service delivery and with it access to shared electronic records. Such records will depend on quality data which must be underpinned by semantic interoperability to realize anticipated benefits [30]. From a business perspective, organizational change science theory specifically highlights the need to ground change programmes in state-of-the-art evidence, and to situate any change in context using co-participatory methods in order to fully understand the nature of the proposed challenges in any planned programme. The Captain's project [11] for example demonstrates how researchers are using innovative co-participatory

S. Das (✉) · P. Hussey
ADAPT Centre, Dublin City University, Dublin, Ireland
e-mail: subhashis.das@dcu.ie; pamela.hussey@dcu.ie

S. Paul et al. (eds.), *Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services*,
EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-77558-2_1

methods in design science through the adoption of living labs and stakeholder communities to advance transformation of systems using technology.

For planned transformation change, the COVID-19 pandemic increases this challenge and presents a global emergency which means more agile and responsive digital resources are required to assist in slowing down the virus transmission. Understanding the complexity and uncertainty of the environment and ensuring that emerging technology is flexible will be key. So too is the strategic fit of the science applied to underpin scheduled programme development and create targeted intervention for meaningful change [22]. In the midst of an ever increasing demographic trend of an ageing population, a pandemic, and a rise in non-communicable diseases, collaboration and cooperation of global research communities is key and has never been more necessary [54]. As a consequence of COVID-19 healthcare professionals are also now increasingly considered vulnerable citizens by policy administrators [56]. We build our systems development life cycle in the midst of a perfect storm and we need robust navigation to deliver sustainable solutions that are effective and provide a good Return On Investment (ROI). Health informatics standards provide guidance to weather this storm and therefore are used to underpin this case study.

1.2 About the CSDM Project

The CeIC[1] in partnership with the Adapt Research Centre[2] in Dublin City University (DCU) are tackling this uncertainty and complexity directly to impact on the aforementioned societal challenges of ageing population and global pandemic. CSDM aligns with the Shanghai Declaration [55] which aims to reorient health and social services to optimize fair access and put people and communities at the centre of policy development. This is particularly important for individuals who are socially underprivileged or with a disability.

Healthy lives and increased wellbeing for people at all ages can be only achieved by promoting health through all of the Sustainable Development Goals (SDGs), and by engaging the whole of society in the health development process including marginalized population groups [53, 55]. In this chapter, we provide insights from a case study in Ireland with our Intellectual Disability (ID) partner services which seeks to address social mobilization with innovative initiatives which help address issues to support our society. We look to service improvement initiatives providing core access to infrastructure for intellectual disability care services delivery. In this chapter, we provide insights on our progress for phase one of this initiative. At the core of the CSDM project, nursing leadership and vision are directing the strategic course and decision making of the research method. This approach is conducted in accordance with international best practice, values and vision of the

[1] https://www.dcu.ie/ceic.

[2] https://www.adaptcentre.ie/.

CSDM team which are decisively rooted in co-participatory care development and delivery processes. Embracing Open Innovation 2.0 principles is key to our work [12, 33]. Cultivating our initial research within the centre, in this chapter we offer a summary of our approach from a computer science perspective. We therefore provide a summary of a specific body of work in progress to deliver a common semantic data model (CSDM) designed using the state-of-the-art health informatics standards from ISO TC215 and CEN TC251 communities (see Table 1.1).

As nursing leadership is at our core, we consider it important to make the case to adopt and position nursing and health informatics theory to lead this interdisciplinary initiative. Drawing on Actor Network Theory (ANT) principles, we consider the nurse (or nurse practitioner) as a central person to communicate and make shared decisions with the patient and their care team to plan and execute a care plan that accounts for all the patient's health conditions. ANT helps us to distinguish how one factor may influence another within any given social process. Focusing on the nurse as an actor and an agent, where the actor is defined as a source of action, we will consider over time the impact and effect on adopted processes in the context of Intellectual disability services focusing case study initially on the development of a COVID-19 app as a non-human agent [5].

1.2.1 Case Study

In our identified Intellectual Disability (ID) partner services the site has a diverse set of care settings including Day service, Residential Services, Community Houses, Independent Living, Host Families—Home Sharing, and an Activity Hub which provides services to the clients and manages a variety of information using paper based and electronic based resources. Managing and providing services for a one-to-one service user and monitoring of patient health status and addressing service user needs has to date worked effectively. With advancement of COVID-19 restrictions implemented through national government guidelines as is the case elsewhere in the world, additional services were identified. Service providers nationally identified the need to enhance the process of work and care flow using digital in health and social care. For example, telehealth increasingly is identified as the primary communication for care delivery. Supporting telehealth and clinical decision making across this diverse healthcare setting is the need for electronic records for integration of care. Such records need to be designed to support interoperability so that data is accessible across and between different systems at the service level. Semantic interoperability allows machine to share, understand, interpret and use data without ambiguity. Lack of semantic interoperability in the healthcare domain is of a great concern, where the main objective is to exchange health-related information with explicit meaning that is shared between different stakeholders or policymakers [4]. The main hurdle in achieving semantic interoperability is the existence of different terminologies used by various coding systems, a lack of correlation among local coding systems and international coding systems which needs to be addressed. An ontology based approach with the standard controlled vocabulary helps tackle the interoperability issues where an explicit formal specification can provide guidance to interact with different systems [31]. A key to understand diverse data and knowledge and big data analysis is ontologies. Ontologies are widely used in biomedical domain and metadata standardization [28], and rigorously support data integration, sharing, re-usability, and computer-based data analysis and automated reasoning. Ontologies are also regarded as the foundation of knowledge representation and knowledge graph, a major field required to support the advancement of artificial intelligence [48]. Making health records meaningful will only be possible if we link the Electronic Health Record (EHR) to an authoritative clinical knowledge and then use natural language in the user interface as suggested by the IMIA working conference on clinical terminology [4]. This enables effective meaning-based retrieval and provides the evidence and associate details to support the adopted specific design approach for CSDM. Using a case study, we can demonstrate and provide guidance to advance knowledge to support eHealth practices.

1.3 Nursing Engagement in System Design

As the largest percentage of the professional workforce, nursing occupies 50% of the total headcount for many countries [57]. There is a wealth of evidence demonstrating the impact of registered nurses on healthcare outcomes for populations across a range of roles particularly missed care [10]. Building on discussion and opinion, the evidence reports that there is also much in the history of nursing which argues the case that the profession needs purposeful and focused engagement [38]. This is conceivable particularly when one considers the significant research completed to identify adverse outcomes on risk which has been published over the past 10 years [2, 34]. Traditionally, nursing has a strong history demonstrating their professions ability to deliver expertise in the realm of systems development and design for health and wellbeing. While the technology may be new, the challenges for generic systems design sadly remain. But without involving interdisciplinary healthcare viewpoint in system design, multi-billion pound losses can happen as evident in Lorenzo healthcare IT system deployed by National Health Service (NHS)-England [51].

Twenty first century healthcare policy advocates advancement of the role of nurse practitioners. The American Medical Informatics Association for example, estimates that as many as 70,000 nursing informatics specialists or analysts will be needed in the next 5 years [49]. From a sociology perspective, Trotter suggests that at its core nursing as a profession can be described as iconic providers of care which provide labour that is gendered and which seeks to support members of society who cannot care for themselves because of age, illness or disability [52]. Interventions such as care coordination with a wide range of people and organizations provide a layered understanding of the broad and variable set of service user needs. Arguing the case that nurses are adequately experienced in advising informing and are therefore experts in transforming the nature of care delivery itself [52]. In this

2 year study, we begin phase one engagement with nursing at the helm directing the development of a case study to advanced strategic tactics of the research.

1.4 Insights on Digital Transformation in Ireland

Ireland is in the process of procuring and deploying a national electronic health record (EHR) programme for the past 5 years. More recent attention focuses on shifting health and social care to the community in the form of a national implementation plan entitled *Sláintecare* [20, 29]. Recent study in the UK shows that Healthcare and social care integration is a key contemporary concept in the service delivery [1]. There is at the time of writing this chapter, however an absence of harmonization of big data which would allow policy analysts and healthcare workers to conduct big data analysis, to reduce costs of treatment, predict outbreaks of epidemics and support disease prevention at the national level. Ireland as a country has invested significantly in registry development in the past; however, data collection within the existing registries cannot support a common semantic data platform and therefore information can only be reported nationally in a fragmented and siloed way [33].

Drawing on relevant evidence we choose to blend Actor–network theory (ANT) principles with an ontology based model and knowledge transfer approach in order for us to be able to publish this research in the open source.[3] This initial work is illustrated through visualization on a graph database platform, GraphDB by Ontotext.[4] Presenting our research development outputs in this way helps build capacity in standards use and deployment through industry. It is important to recognize that this approach is not suitable for use with a relational database which is the more conventional approach in healthcare for data storage for secondary care delivery.

We also argue the case that future models of care aligned with our national and global policy will need to be linked with the Internet of Things (IoT) to advance self-management support and will require access to data across one to many healthcare settings. This can be only achieved if we use an ontology based model which is relying on linked data approach and strategically fits better for the digital patient's future requirements. According to market research there is currently in 2020 3.68 million smartphone users in Ireland.[5] On July 6th, 2020 the public health services app for tracing citizens with COVID-19 was launched in Ireland and passed one million registrations within 48 h of its launch [43]. In this context, we consider it important to focus on a case study approach to help us to understand and assist in the development of a model with intellectual disability services. A model where

[3] https://bioportal.bioontology.org/ontologies/CONTSONTO.

[4] https://www.ontotext.com/products/graphdb/.

[5] https://www.statista.com/statistics/494649/smartphone-users-in-ireland/.

the terminology is motivated by case study scenarios that focus on clinical need and capable of supporting future service improvement. Phase one of this project reported here provides deliverables which offers the team a guidance so that we are able to understand requirements as well as identify limitations which helps us during modelling and future implementation of the systems development lifecycle.

1.5 Summary of Case Study Methodology Phase One

Phase one of the CSDM research project aims to set in place a framework to develop a system architecture that will be useful for other application scenarios in the future. Preliminary output of our research development work is available in open source domain.[6] Drawing on specific evidence such as the WHO policy on patient centred integrated care [54] and the work of Blobel [6–9], we also contribute to European Committee for Standardization (CEN) and ISO community working groups on health informatics [15, 35]. The nursing led team established a working group to define a use case and create and deploy an extension of Semantic Sensor Network (SSN) ontology [27] and Fast Healthcare Interoperability Resources (FHIR) [32] application programming interface (API) with an industry partner entitled Davra[7] in quarter 2 of 2020. Through initial analysis for sustainable health systems access, the following information was collated and critiqued.

1.5.1 Initial Requirements Gathering and Needs Analysis

Guided by Blobels' systems approach to interoperability [6] in combination with OntoClean methodology [25], the team created a formalized ontology and defined an information model to underpin the project in phase one. The information model for data, information, and action was adopted in partnership with focused discussion groups and an advisory board was identified with dedicated implementation team. Included in the advisory board are two Fulbright academic scholars with a background in doctoral nursing practice from the USA who provide insights into the project action and data management plan from a global perspective.

Some early key milestones from preliminary discussions included the need to develop a common business case underpinned by two communicating and cooperating principles. In our case these identified principals were *quality* and *relevance* to inform contributions and deliverables. The first principal was to ensure that identified contributions be as informative as is required for the purpose of exchange of semantic data. The second principal was to ensure that the contributions are

[6] https://opensource.org/.

[7] https://davra.com/.

focused on relevance of identified requirements with a view to avoiding obscurity of expression in community services. The team recognized the nature of this work in the context of a COVID-19 pandemic and considered the information cycle process in a phased manner to deploy existing or emerging knowledge, skills and capabilities within the team. Agreed agendas included building capacity for shared communications and dissemination of findings as part of the communication and cooperation process across the CSDM project. A critical review of the evidence to informed decision making on the case study provided the flowing information and action steps as detailed in the following Sect. 1.2.

1.5.1.1 Evidence and Action Steps

The ISO Reference Model for Open Distributed Processing (ODP) was used to analyse the emerging tools, resources and standards published on the various interoperability levels using core evidence from Blobel [6–9]. This provided a rigorous structure and pathway to build capacity and progress with the project specified requirements to deliver on the following goals in phase one. These standards are outlined in Table 1.1 using the various viewpoints for ODP with the exception of the technical viewpoint scheduled for review later in this case study development lifecycle. Goals and scoping review were considered both informally (Social participatory engagement processes) and formally (Informatics technical implementation processes).

Informal: To develop and report with advisory boards and dedicated implementation team a proof of concept research study for assessment of service users within a specific intellectual disability (ID) service. This person-centred initiative focuses on COVID-19, Referral and Healthy ageing initiatives. Design of an information data plan will be underpinned with related health policy—*Sláintecare*, eHealth Ireland and ID national policy. *Formal*: Drawing on ANT and co-participatory methods develop phase one of a Common Semantic Data Model (CSDM) for an emerging interoperable framework capable of accommodating different schemas for knowledge transfer on health informatics standards for digital health transformation with ID services.

1.5.2 Scoping Review and Mapping Exercise

In order to define the Interoperable-Integrated care Service Architecture (IISArc), a qualitative and quantitative analysis was carried out to investigate the correlation between different actors involved in the residential care service. We also investigated the correlation between several smart devices including IoT based services for residential care and the corresponding needs of the service users with COVID-19 through detailed clinical modelling and testing in a simulated environment [41].

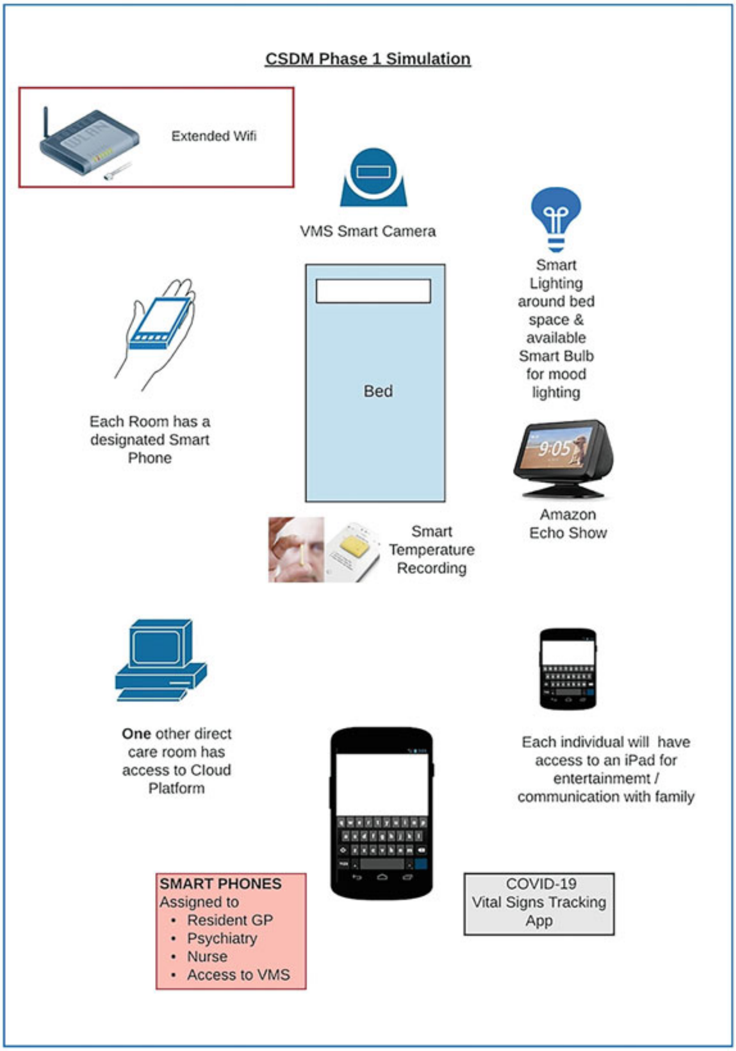


Fig. 1.1 Infrastructure view of a stimulated residential care unit

Figure 1.1 depicted typical arrangement of medical device, IoT, smartphone to collect patient observation data within our identified residential care unit.

This initiative was then tested in a day service with nurse practitioner leading a simulated exercise to test clinical workflow and impact of digital devices in context. As individuals with Intellectual disability (ID) are considered a vulnerable population group and COVID-19 cocooning was in process, for preliminary work conducted in phase one we opted not to directly involve the service users in the initial design process. The co-participatory group opted to work with nurse

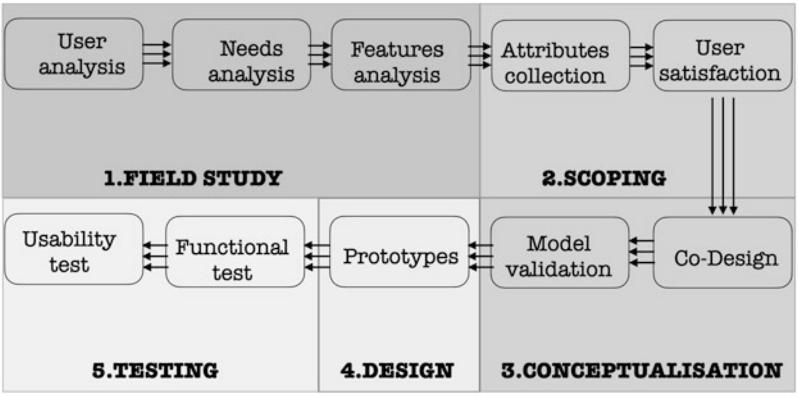


Fig. 1.2 Process flow schema adopted for the IISArc project

practitioners and healthcare professionals and experienced carers who provided detail for the mapping of the service. The plan being to provide a demonstrator to service users for review and comment and additional development as required once ethical approval has been obtained. Overall on-going design and workflow for IISArc, shown in Fig. 1.2 and comprise of the following five steps: (1) field study, (2) scoping, (3) conceptualization, (4) design and (5) testing.

Field Study An ethics proposal was drafted for the clinical engagement phase of the project. As no personal or identifiable records were to be used for the initial scoping review and mapping exercise in phase one, there was no anticipated risk associated with the study. Simulated data was created to demonstrate the testing and demonstration of the prototype. Primary users for phase one identified in the case are nurses and carers. Secondary users identified are the management team in the service, the individual clients with intellectual disability (ID) and the Industry partner Davra. User centric-design has been applied to develop the users' needs analysis based on evidence on COVID-19. It is anticipated that the project scope for phase one which has been developed in the ID services may be useful to address older person needs in the future. This aligns well with the overarching mission of the research team to ensure an adaptive, usable and understandable sustainable product is delivered which may be adapted for other domains in the future.

In Fig. 1.3, we will describe process, domain and context of use. We illustrate a summary view of how we are defining requirements of the information layer for semantic interoperability in the identified domain using a phased and step by step cyclical process. The context of use is defined with practitioners who consider workflow and data requirements for assessment of a COVID-19 service user. We consider a deteriorating service user who requires transfer from a social care residential unit to an acute care provider. Core data for transfer in line with national guidelines and specific care requirements for the intellectual disability domain are taken into consideration [30]. The user needs and feature analysis

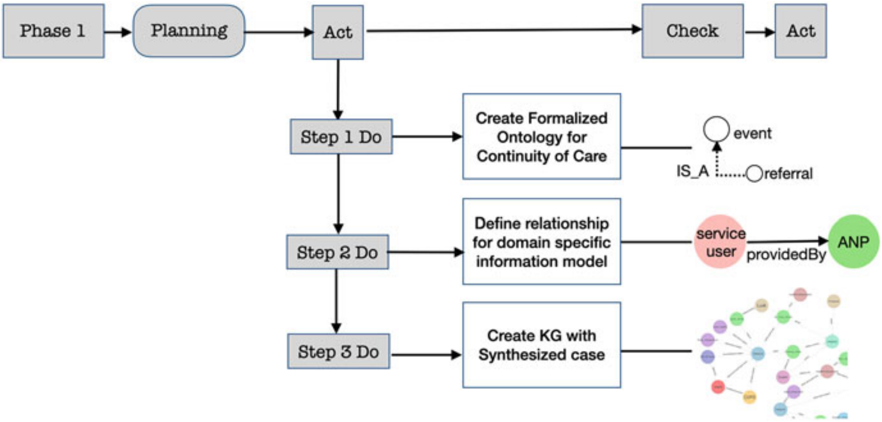


Fig. 1.3 Plan–Do–Study–Act adapted for CSDM

include focus group interviews and documentary analysis completed by health care practitioners which are conducted as phase one. This is followed by a series of steps which confirm with the practitioners the specific relationships between the various resource requirements for data transfer across different care settings. We consider the focus to be on assessment of service needs for an individual service user and referral of key information for transfer with the service user to support safe referral of care. The final step in this phase of needs analysis and scoping process relates to the development of the knowledge graph (KG) using synthesized cases for demonstration to the wider stakeholder group through planned dedicated workshops. At each step of the CSDM project, there is a check action to review and confirm that CSDM resource is aligning with the domain and context of use.

1.5.2.1 Informal Scoping

Initial attributes collection as part of a scoping phase was completed outside of the service with the experienced nursing group. The main objective being to understand which attributes aka data fields are needed for our initial conceptual model development. Key questions which we have used within the focus group discussions with service providers relate to the following: What are the most relevant and agreed requirements considered essential for day to day patient observation data collection? These reported data attributes are currently reported to be in form of paper or excel spreadsheets from existing information system deployed in the residential care facility. Step one and step two of the user and needs analysis was completed to identify the data attributes to create an initial conceptual model. An example of the data attributes identified and related to the COVID-19 app is available from FHIR [32]. For example, temperature and breathing are included in the FHIR observation detail and available in Sect. 1.5.3.3.

In adopting a hybrid approach to tackle the complexity of health and social care delivery two specific standards Fast Healthcare Interoperability Resources (FHIR) and Observational Medical Outcomes Partnership (OMOP) [32, 44] were used. FHIR has well-defined data structure (aka Resource) to capture patient observation data as well as other components involved in a care setting such as patient, diagnostic report, medication statement, etc. Detailed explanation on how this case study use FHIR is available in Sect. 1.5.3. OMOP was used to guide the dataset mapping coming from across care settings.

Design Principles Agreed key design principles adopted by CSDM are in accordance with core principles of health interoperability [4]. The project team will accommodate and store observational data to facilitate ongoing research which is:

- Suitability for purpose
- Data protection
- Rationale for domains
- Standardized Vocabularies
- Reuse of existing vocabularies
- Maintaining source codes

Initial scoping review of health informatics standards and resources used are presented in Table 1.1 under the Reference Model of Open Distributed Processing (RM-ODP) structure. RM-ODP is an architectural standard developed in the mid-1990s following 10 years of development work through the International Standards Community on architecture [39]. This standard explains the inter relations between the various viewpoints as the enterprise viewpoint linking with the system designer role, the information viewpoint which informs the information architect role, and the component developer which uses the computer viewpoint and the systems implementer role which uses the engineering viewpoint supported by the systems administrator role in the technology viewpoint. The technical viewpoint is not included in Table 1.1 as this review is scheduled for later in the project lifecycle when tested in the service context [6, 39].

1.5.2.2 Formal Scoping

Identification of the formal scoping and mapping exercise involved the analysis of the *Conceptualisation* model. This action step commenced with a focus on the co-design and model validation within the core research team and four focused co-participatory group discussions. These discussions were considered preliminary because of COVID-19; access to the services was delayed. The team opted to use ISO 22272 a Health Informatics Standard Methodology for analysis of business and information needs of health enterprises to support standards based architectures. These initial discussions explored the project from four specific perspectives to consider the *Business Impact Analysis*, *Concept Analysis*, *Information Analysis* and *Future Model Development*. In the context of CSDM this phase one analysis

Table 1.1 Summary table of standard and techniques use in our project

Standards and techniques	Category	Framework development utility
ISO 22272	Enterprise	Used to define Business mission aims and objectives, conceptual model and capacity building
Contsys ISO 13940 https://contsys.org/	Computer	Overarching ontology framework designed in OWL as a formalized web ontology (publish on May 2020) https://contsys.org/pages/Guest%20blog/FormalOntology
HL7 FHIR v.4 https://hl7.org/FHIR/	Computer	Used to specify attribute structure of Observation, Medication Statement, Medication Administration, Medication Diagnostic, Report, Patient
EHRcom ISO 13606	Engineering	Reviewed in terms of harmonization of HI standards to describe distribution of objects and applications across platforms
IHE CCR https://www.astm.org/Standards/E2369.htm	Information	Provides a patient health summary standard which facilitates timely and focused transmission of information across and between health professionals involved in the delivery of patient's care
Snomed-CT ICNP	Information	Formalized reference terminologies used for semantic mapping of concepts in health care systems
Protégé https://protege.stanford.edu	Engineering	A free, open source ontology editor and framework for building intelligent systems
Karma https://usc-isi-i2.github.io/karma/	Engineering	A data integration tool and aligning with Interoperability level-1 Technical automation. Semi automation of semantic mapping is important so that the semantic detail used is approved and validated by nurses
GraphDB http://graphdb.ontotext.com	Engineering	Used as a specification with common functions providing a technology-neutral architectural framework and database support

on the conceptual model refers to identifying the core design requirements for the ontological model which enables semantic interoperability [46]. This formal scoping also included a number of presentations of the evidence selected to support the CSDM project and a detailed explanation on how ontologies can support future models of integrated care development. A defined ontology can be viewed as a declarative model of a domain that defines and represents the concepts existing in that domain, their attributes and the relationships between them. It is typically represented as a knowledge base which then becomes available to applications that need to use and/or share the knowledge of the identified domain. In addition the emerging conceptual model used Web Ontology Language (OWL). In the domain of health informatics, an ontology [23] can be described as a formal description of a health-related domain. OWL offers a logic-based semantic markup language for publishing and sharing ontologies on the World Wide Web [45]. OWL 2 was used [19] and found useful for our project particularly in demonstrating specific detail of the case study with key stakeholders. OWL 2 was selected as it supports both qualified and unqualified cardinality restrictions. Its datatypes include various kinds of numbers, adding support for a wider range of XML Schema Datatypes (double, float, decimal, positive integer, etc.) and providing its own datatypes.

This was considered a useful resource to linked open data community. In addition to developing a preliminary conceptual model in OWL, the development team also conducted some terminology mapping as part of the *model validation process*. Specifically focusing on changing the values found in the source datasets into the values defined by the standard organization resources adopted in the project lifecycle. Typically data elements are stored in a database which consists of values such as a word, number or a date, a text string or a code. Each database uses values that are defined by the database designer, in this case an intellectual disability (ID) service provider. The defined set of codes can be called a reference terminology. In healthcare within Ireland existing widely used and adopted terminologies include International Classification of Diseases 10th revision (ICD-10), and more recently Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) for diagnoses and intervention mapping. The CSDM development team considers it important to recognize that many databases contain values that are unique to their local environment [3, 4]; therefore, differences in codes and values can lead to inconsistencies where data are combined across many different local systems resulting in heterogeneous data collections. This reflects the challenges ahead for national programmes in adoption of national electronic health records (EHR) and why the European Union (EU) is focusing on ICT standardization in the EU with programmes such as the Elite S Scholarship [14] which is funding this research. What is required for existing data banks is cleaning of data for terminology mapping in order to convert local terms into standard terms that are used across all data partners or care centres in a healthcare network. Figure 1.4 provides a summary of the tooling used in this phase of the CSDM which addresses this issue by adopting language independent ontological model which helps to integrate data based on their semantic meaning rather terms.

1.5.3 Technical Implementation

Dissemination of information on the CSDM is provided on a phased basis with the ISO and CEN Community, for example publication of formal ontology for community of care is available with a supporting blog post on the Contsys Website.[8] For technical implementationFor technical implementation depicted In Fig. 1.4, provides a summary of CSDM intended implementation pipeline which includes tools and techniques on how we executed our work to date.

As we await ethical approval for data collection, we will then collect more data from different healthcare data sources through our identified service. We anticipate

[8] https://contsys.org/pages/Guest%20blog/FormalOntology.

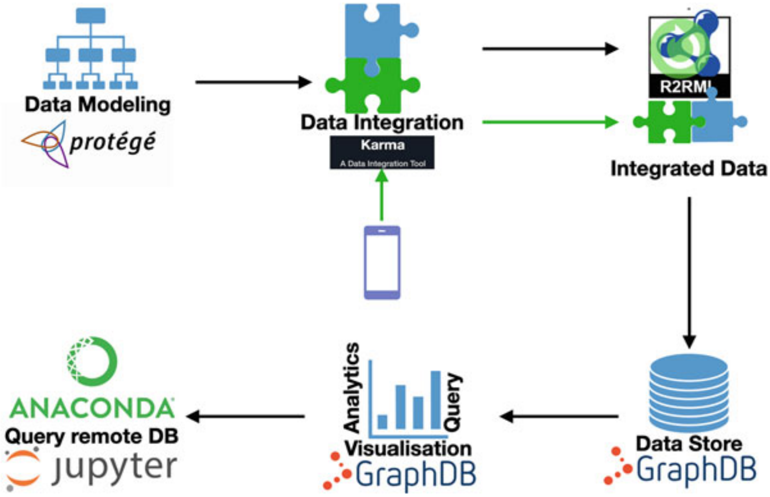


Fig. 1.4 Implementation pipeline used for CSDM

that this will be possible through planned fieldwork, such as workshop and survey activity as this approach has worked well for preliminary work conducted in phase one of this study. As required the development team will then set about cleaning the data collated using some software such as ontoRefin, which is a data transformation tool, based on OpenRefine[9] and integrated in the GraphDB Workbench [21]. Protégé [42] is a free, open source ontology editor and a knowledge management system, which will be used for designing and editing IISArc schema. Protégé provides a user-friendly graphical interface to define ontologies. In particular for defining the terminology and schema mapping Cellfie [36], a Protégé desktop plugin will be used for importing spreadsheet data into OWL ontologies for data integration tasks. As the research programme grows for large dataset we will use KARMA, for integration [26]. KARMA is an open source tool which enables to integrate data from different sources like XML, CSV, text files, web Application Programming Interface (API's). KARMA also generates RDB to RDF Mapping Language (R2RML) mapping file which can be reused again and again in case of feeding models with new data.

Other features under consideration for System Architecture of CSDM are as depicted in Fig. 1.5. We aim to connect with an existing local IT system called Clinical Information System (CIS); this is a particular in-house system of the service organization which has been in place for a number of years. The Geographical Information System (GIS) applications could be used to collect and locate service user and staff details including emergency management planning, and IoT devices to monitor service user's health and wellbeing. The local IT information system CIS

[9]https://openrefine.org/.

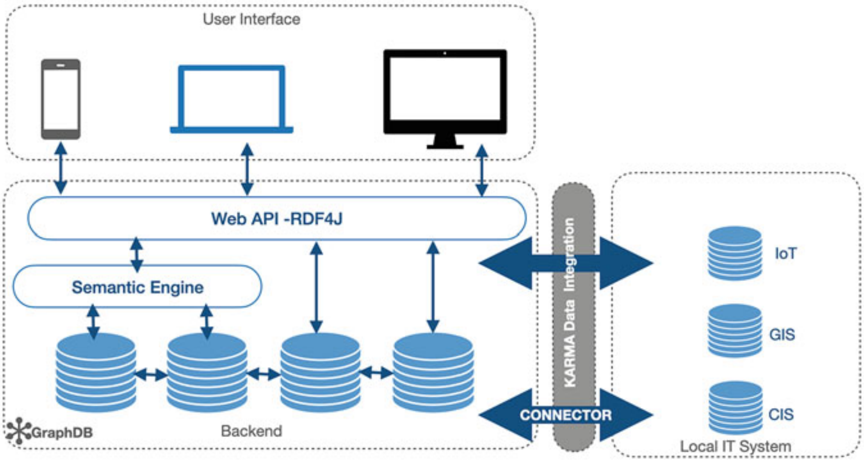


Fig. 1.5 System architecture of CSDM

could be connected with the GraphDB via a connector while KARMA could be used to harmonize the data. This could then facilitate access by staff and service users to information via organizational laptop, desktop or other mobile devices which are connected with the system via web API.

1.5.3.1 Data Modelling

As it is our first phase of the project, we mainly focus on data modelling part to generate initial IISArc ontological schema. Full formal ontology for continuity of care comprises around 138 classes.[10] Out of that main classes needed to capture our scenario are rather small. They are:

- *Healthcare actor* organization or person participating in healthcare. The involvement of the healthcare actor will be either direct (for example, the actual provision of care) or indirect (for example, at organizational level).
- *Subject of care* healthcare actor with a person role; who seeks to receive, is receiving, or has received healthcare. Synonym: subject of healthcare; service user; patient; client.
- *Healthcare profession* having a healthcare professional entitlement recognized in a given jurisdiction. The healthcare professional entitlement entitles a healthcare professional to provide healthcare independent of a role in a healthcare organization.

[10] see https://bioportal.bioontology.org/ontologies/CONTSONTO.

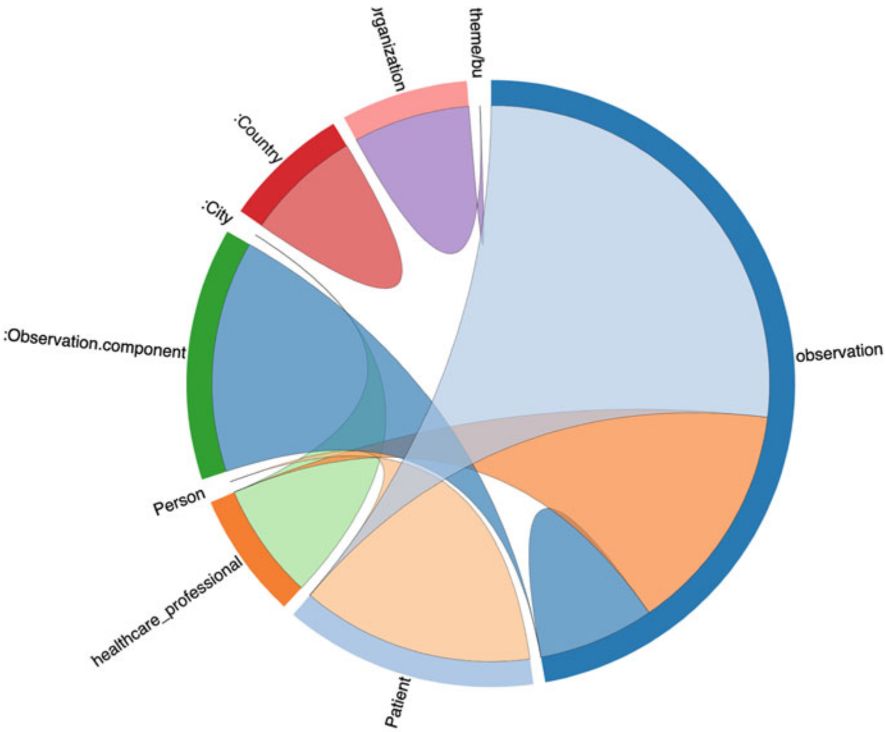


Fig. 1.6 Class relationship with Observation

- *Observation* Observations are a central element in healthcare, used to support diagnosis, monitor progress, determine baselines and patterns and even capture demographic characteristics. We reuse HL7 FHIR observation [32] Resource in our ontology which provides the details and data structure for capturing patient daily observation. Class relationship with *observation* class is depicted in Fig. 1.6. It represents class relationships among the nine classes which contain data and how they are interrelated with each other. We use Ontotext GraphDB [21] to generate this class diagram. The Class relationships diagram is based on the real statements (i.e. instance level) between classes and not on the ontology schema.

1.5.3.2 Ontology Alignment

Top-level ontologies provide domain-independent conceptualization, relations and axioms (e.g. categories like Event, Mental Object, Quality, etc.) in order to standardize the ontology creation by providing guideline based on a rigorous ontological commitment. Therefore, building the continuity of care ontology under

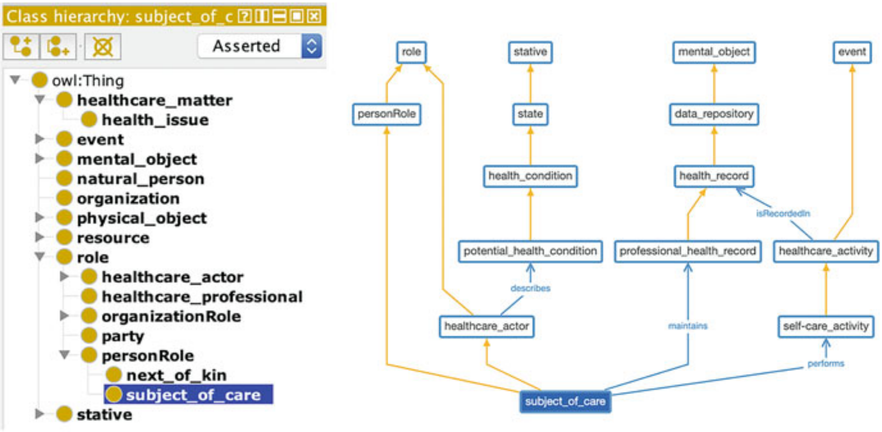


Fig. 1.7 Formal Ontology of continuity of care developed and visualized using Protégé ontology editor

a top-level ontology should contribute to improve its semantic interoperability with other existing ontologies. In our work, we use the top-level ontology Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [24] as a middle-out solution between degree of formalization and complexity, contributing to an effective practical solution. In spite of the benefit of top-level ontologies, their alignment and use is not trivial and requires some expert effort. The EU project Advancing Clinico-Genomic Trials (ACGT) [50] as well as other healthcare projects emphasize on need and benefit from top-level alignment.

DOLCE ontology classes such as stative, mental object, event, natural person and physical object preoccupied class hierarchy of our ontology. This helps to aggregate similar kind of classes of continuity care under one super class. The European Federation for Medical Informatics (EFMI) also emphasises the need for ontological scrutiny for continuity of care [40] however it is interesting to note we found no further progress in our review in the past decade. Our formal model tries to resolve those ontological issues as well with our collaborative work with the existing ISO 13940 community. Figure 1.7 depicted fragment of our formal OWL ontology in Protégé user interface.

1.5.3.3 Formal Model

In the research completed to date the expressiveness of our ontology model is *ALCHQ(D)* as per description logic (DL) scale. We opted not to exploit full power of DL-Full as supported by OWL language, rather we opted to use simple rule so that it is compatible with a GraphDB rule engine. This approach facilitates quick query execution time as we are cognisant that our model will have potential to expand exponentially in the future. In this section we provided details on the data structure of the class *observation* and property *person.gender* in Resource Description Framework (RDF) Turtle syntax. The observation class defined by FHIR Observation accommodates reuse of properties in order to facilitate our model for interoperability. This approach aligns with other EHR models using FHIR based specifications and supports cross border studies on intellectual disability (ID) clients in the future. Observation captures measurements and simple assertions made about a patient, device or other subject. The subject of observation is *Patient* class. The performer of observation is *healthcare_professional* class. Datatype restriction is encapsulated under property *person.gender*, where the data provider has to choose among gender value "Male", "Female" or "Transsexual" as it is a mandatory and important information needed by the service provider. It cannot be left blank and reasoner will able to detect in case of wrong information inserted in the system. The following excerpt demonstrates an example of this observation class detail in RDF turtle syntax.

observation(class)

```
http://ontologies.dcu_smh.ie/CeIC-SMH#Observation
EWS:Observation rdf:type owl:Class ;
    rdfs:subClassOf [ rdf:type owl:Restriction;
    owl:onProperty observation-definitions:Observation.performer;
    owl:someValuesFrom EWS:healthcare_professional],
    [rdf:type owl:Restriction;
    owl:onProperty observation-definitions:Observation.subject;
    owl:someValuesFrom EWS:Patient];
oboInOwl:hasDbXref
"https://www.hl7.org/fhir/observation-definitions.html
#Observation"^^xsd:anyURI;
rdfs:comment "Measurements and simple assertions made about
a patient,device or other subject."@en.
```

person.gender(property)

```
https://www.hl7.org/fhir/person-definitions.html#Person.gender
person-definitions:Person.gender rdf:type owl:DatatypeProperty ;
  rdfs:range [ rdf:type rdfs:Datatype ;
  owl:unionOf ( [ rdf:type rdfs:Datatype ;
  owl:oneOf [ rdf:type rdf:List ;
  rdf:first "Female"^^xsd:string ;
  rdf:rest rdf:nil]]
  [ rdf:type rdfs:Datatype ;
  owl:oneOf [ rdf:type rdf:List ;
  rdf:first "Male"^^xsd:string ;
  rdf:rest rdf:nil]]
[ rdf:type rdfs:Datatype ;
  owl:oneOf [ rdf:type rdf:List ;
  rdf:first "Transsexual"^^xsd:string ;
  rdf:rest rdf:nil]])] ;
  rdfs:comment "The gender might not match the biological sex
  as determined by genetics,or preferred identification.
  Note that there are other possibilities than M and F".
```

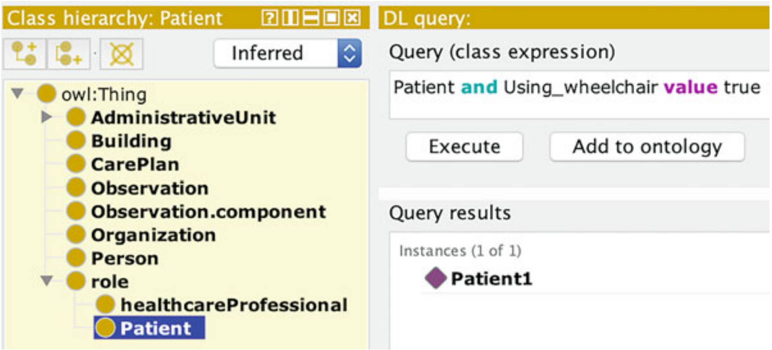


Fig. 1.8 DL Query execution with HermiT reasoner

1.5.3.4 Inference

OWL contain schema information in addition to links between different classes. This additional information and rules allow users to perform reasoning on the knowledge bases in order to infer new knowledge and expand existing knowledge. The core of the inference process is to continuously apply schema-related rules on the input data to infer new facts. This process was considered helpful in this case study for deriving new knowledge and for detecting inconsistencies. Example of automated inference using HermiT [47] reasoner is depicted in Fig. 1.8.

1.5.3.5 Knowledge Graph

Knowledge graph (KG) can be briefly explained as a graph with interconnected entities. KG became popular in early 2012 when Google started showing their search result as a Knowledge Graph which appeared on the right side of the page of your search result [13]. KG feature is available in all Graph database technology. In this case study, we obtain our initial IISArc KG using Ontotext GraphDB as shown in Fig. 1.9.

We can see how observation, patient and healthcare professional are connected. Here node represents entity and edge represents relation. Node with same colour means those entities belong to the same ontology class. On the right side of the illustration, we can see information related to particular node, in this case for *Diastolic blood pressure*. The user can click on different nodes to know detailed information related to that entity. It is a very intuitive way of visualization and also very helpful for healthcare professional or even healthcare management group to show workflow within identified residential care facility. Our initial interview with service providers indicates their satisfaction on our preliminary result. The next challenge is to formally evaluate the resource once additional data is entered from different settings.

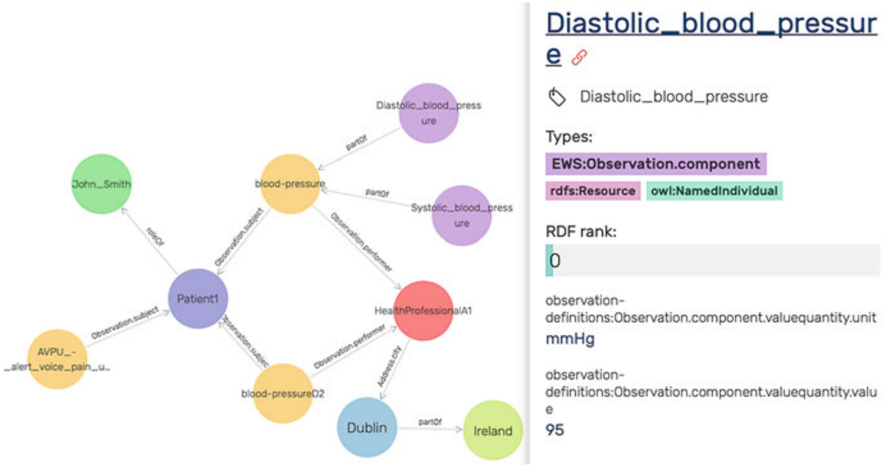


Fig. 1.9 Initial Knowledge Graph of IISArc visualized using GraphDB

1.6 Discussion

Considering the scope of the project and the need to identify alternative care delivery mechanisms in the era of COVID-19, this nurse led initiative recognized the need to deal with the challenges at the organizational level and devise innovative approaches to progress the CSDM project. Nurses are critical providers of care and therefore contextually understand the organization and what is required to achieve intended goals and service user outcome. The COVID-19 pandemic intensifies not only the economic needs of service provision but also the social needs of service users, and research needs to be cognisant of this when designing systems for new models of care delivery. Focused coordination and social distancing guidance on key requirements of COVID-19 can generate as many problems as they are designed to solve. It is therefore critical to approach the project with an interdisciplinary focus using a layered understanding of the existing and potential problems created by changing care delivery mechanisms. Simulation of planned transitions carries low risk and can be completed ahead of a study which requires ethical approval. This preliminary work has provided insights for the CSDM project and provides early indicator on the practical application and use of standards and technical reports for implementation of a common semantic data model for a complex healthcare network. As recommended by OECD [57], even though medicine has been successful at treating disease continual initiatives for service improvement must include assessment of the impact treatments have on people’s lives. This makes outcomes valued by patients a key indicator of success.

1.7 Conclusion and Future Work

The European Union, aware of the poor uptake of health informatics standards, and associated fragmentation of health systems delivery has introduced the Elite S scholarship programme for future leadership in ICT standardization in Europe. In DCU through CeIC and Adapt Research Centers, researchers are working in partnership to accelerate the uptake and use of health informatics standards in Ireland. The COVID-19 pandemic has presented many challenges globally, and health researchers and policy analysts are reporting worrying results on predictive models for 2020–2021. Health and social care professionals are now classified as vulnerable populations by WHO [57]. It is critically important that such vulnerable employees are not unduly placed at risk, and in their respective practices that their time is not unduly wasted doing menial and repetitive tasks that can better be achieved using alternative and more efficient practices. Digital health systems and services can potentially provide some leverage in addressing some of the more outdated and repetitive routines that take up a significant HCP time. Co-participatory methods to advance appropriate design science of technology in context can potentially impact on service providers. Such systems once deployed can leave more time for important interventions that can address some of the more pressing needs in direct care provision. In addition to the longer term impact by providing data analytic and visualization for population outcome monitoring and policy planning.

During phase one development the CSDM team has identified opportunity in the current working processes with potential for service improvement. The next step is showcasing the benefit of using integrated-interoperable care service architecture (IISArc) to the wider audience. In this way the benefit of adopting graph database as part of a data storage and long term benefit to manage and retrieve service information can be realized. Additional benefit include but not limited to data harmonization and discovering new knowledge to inform targeted interventions. The second phase of development will test CSDM application from user satisfaction point of view using user experience (UX) dimension as proposed by Laugwitz [37] and as we implemented in our previous work on Semantic User Interface (SemUI) [18]. The questionnaire will be designed to understand different UX dimensions [37] along with the specific traits of *User Interface*. These UX dimensions perform a thorough assessment of the product using 6 scales with 26 terms. These scales were: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty. Final step will be to deploy our system on a cloud based server with the help of our industry partner Davra in order for us to conduct performance testing of the CSDM platform.

Acknowledgments This research has received funding from the European Union's Horizon 2020 research and innovation programme under the ELITE-S Marie Skłodowska-Curie grant agreement No. 801522, by Science Foundation Ireland and co-funded by the European Regional Development Fund through the ADAPT Centre for Digital Content Technology grant number 13/RC/2106 and DAVRA Networks.

References

1. A. Altabaibeh, K.A. Caldwell, M.A. Volante, Tracing healthcare organisation integration in the UK using actor–network theory. J. Health Organ. Manag. **34**(2), 192–206 (2020)
2. D. Ausserhofer, B. Zander, R. Busse, M. Schubert, S. De Geest, A.M. Rafferty, J. Ball, A. Scott, J. Kinnunen, M. Heinen, et al. Prevalence, patterns and predictors of nursing care left undone in European hospitals: results from the multicountry cross-sectional rn4cast study. BMJ Qual. Saf. **23**(2), 126–135 (2014)
3. G. Bella, L. Elliot, S. Das, S. Pavis, E. Turra, D. Robertson, F. Giunchiglia, Cross-border medical research using multi-layered and distributed knowledge, in *ECAI* (2020), pp. 1–8
4. T. Benson, G. Grieve, *Principles of health interoperability: SNOMED CT, HL7 and FHIR* (Springer, Berlin, 2016)
5. M. Berg, Patient care information systems and health care work: a sociotechnical approach. Int. J. Med. Inform. **55**(2), 87–101 (1999)
6. B. Blobel, Interoperable EHR systems–challenges, standards and solutions. Eur. J. Biomed. Inform. **14**(2), 10–19 (2018)
7. B. Blobel, Challenges and solutions for designing and managing pHealth ecosystems. Front. Med. **6**, 83 (2019)
8. B. Blobel, M. Giacomini, Interoperability is more than just technology. Eur. J. Biomed. Inform. **12**, en1–en2 (2016)
9. B. Blobel, M. Giacomini, *pHealth 2019: Proceedings of the 16th International Conference on Wearable Micro and Nano Technologies for Personalized Health 10–12 June 2019, Genoa, Italy*, vol. 261 (IOS Press, Amsterdam, 2019)
10. M. Butler, T.J. Schultz, P. Halligan, A. Sheridan, L. Kinsman, T. Rotter, J. Beaumier, R.G. Kelly, J. Drennan, Hospital nurse-staffing models and patient-and staff-related outcomes. Cochrane Database Syst. Rev. **1**(4), 1–63 (2019)
11. Captain-EU. Coach assistant via projected and tangible interface: smart home assistant for older adults (2018). https://www.captain-eu.org/project-overview
12. M. Curley, B. Salmelin, *Open Innovation 2.0: The New Mode of Digital Innovation for Prosperity and Sustainability* (Springer, New York, 2017)
13. J.S. Eder, Knowledge graph based search system, June 21 2012. US Patent App. 13/404,109
14. ELITE-S. Future leadership in ICT and standardization in Europe (2019). https://elite-fellowship.eu/
15. European Committee for Standardization. Cen/tc 251 - health informatics (2019). https://standards.cen.eu/dyn/www/f?p=204:7:0::::FSP_ORG_ID:6232&cs=18CA078392807EDD402B798AAEF1644E1
16. L. Fæste, M. Reeves, K. Whitaker, Science of organizational change, winning the '20s, Boston consulting group (2019). https://www.bcg.com/publications/2019/science-organizational-change
17. E. Fry, F. Schulte, Death by a thousand clicks: where electronic health records went wrong. Fortune Mag., **3**, 18 (2019)
18. F. Giunchiglia, S.R. Ojha, S. Das, Semui: A knowledge driven visualization of diversified data, in *2017 IEEE 11th International Conference on Semantic Computing (ICSC)* (IEEE, Piscataway, 2017), pp. 234–241
19. C. Golbreich, E.K. Wallace, P.F. Patel-Schneider, Owl 2 web ontology language new features and rationale. *W3C working draft, W3C (June 2009)*. http://www.w3.org/TR/2009/WD-owl2-new-features-20090611 (2009)
20. Government of Ireland. Sláintecare implementation strategy (2019). https://www.gov.ie/en/publication/6011cf-slaintecare-implementation-strategy/#
21. GraphDB. Graphdb workbench (2019). https://graphdb.ontotext.com/documentation/standard/workbench.html
22. T. Greenhalgh, J. Wherton, C. Papoutsi, J. Lynch, G. Hughes, S. Hinder, N. Fahy, R. Procter, S. Shaw, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption,

abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research* **19**(11), e367 (2017)

23. T.R. Gruber, A translation approach to portable ontology specifications. Knowl. Acquis. **5**(2), 199–220 (1993)
24. N. Guarino, *Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6–8, Trento, Italy*, vol. 46 (IOS Press, Amsterdam, 1998)
25. N. Guarino, C.A. Welty, An overview of OntoClean, in *Handbook on Ontologies* (Springer, Heidelberg, 2004), pp. 151–171
26. S. Gupta, P. Szekely, C.A. Knoblock, A. Goel, M. Taheriyan, M. Muslea, Karma: A system for mapping structured sources into the semantic web, in *Extended Semantic Web Conference* (Springer, Berlin, 2012), pp. 430–434
27. A. Haller, K. Janowicz, S.J.D. Cox, M. Lefrançois, K. Taylor, D. Le Phuoc, J. Lieberman, R. García-Castro, R. Atkinson, C. Stadler, The SOSA/SSN ontology: a joint WEC and OGC standard specifying the semantics of sensors observations actuation and sampling, in *Semantic Web*, vol. 1 (IOS Press, Amsterdam, 2018), pp. 1–19
28. J. Hartmann, Y. Sure, P. Haase, R. Palma, M. Suarez-Figueroa, OMV–ontology metadata vocabulary, in *ISWC*, vol. 3729 (2005)
29. Health Service Executive. ehealth Ireland (2015). https://www.ehealthireland.ie/knowledge-information-plan/ehealth-strategy-for-ireland.pdf
30. Health Service Executive. How can telehealth best support HSCP's response to the COVID-19 patient? (2020). https://hselibrary.ie/how-can-telehealth-best-support-hscps-response-to-the-covid-19-patient/
31. G. Héja, G. Surján, P. Varga, Ontological analysis of SNOMED CT. BMC Med. Inform. Decis. Mak. **8**(1), S8 (2008)
32. HL7. FHIR release 4 (2019). https://www.hl7.org/fhir/
33. P. Hussey, K. McGlinn, The role of academia in reorientation models of care–insights on ehealth. Informatics **6**(3), 37 (2019)
34. ICN. Nursing now campaign (2020). https://www.icn.ch/news/international-council-nurses-and-nursing-now-welcome-2020-international-year-nurse-and-midwife
35. ISO. ISO/TC 215 health informatics (2019). https://www.iso.org/committee/54960.html
36. H. Josef, Cellfie-plugin user's guide (2016). https://github.com/protegeproject/cellfie-plugin
37. B. Laugwitz, T. Held, M. Schrepp, Construction and evaluation of a user experience questionnaire, in *Symposium of the Austrian HCI and Usability Engineering Group* (Springer, Berlin, 2008), pp. 63–76
38. M. Laurant, M. van der Biezen, N. Wijers, K. Watananirun, E. Kontopantelis, A.J.A.H. van Vught, Nurses as substitutes for doctors in primary care. Cochrane Database Syst. Rev. **1**(7), 1–26 (2018)
39. P.F. Linington, RM-ODP: the architecture, in *Open Distributed Processing* (Springer, Berlin, 1995), pp. 15–33
40. C. Martínez-Costa, S. Kay, N. Oughtibridge, S. Schulz, Contsys under ontological scrutiny, in *MIE* (2015), pp. 999
41. G. Mincolelli, S. Imbesi, G.A. Giacobone, M. Marchi, Internet of things and elderly: Quantitative and qualitative benchmarking of smart objects, in *International Conference on Applied Human Factors and Ergonomics* (Springer, Berlin, 2018), pp. 335–345
42. N.F. Noy, M. Crubézy, R.W. Fergerson, H. Knublauch, S.W. Tu, J. Vendetti, M.A. Musen, Protégé-2000: an open-source ontology-development and knowledge-acquisition environment, in *AMIA... Annual Symposium proceedings. AMIA Symposium* (2003), pp. 953–953
43. C. O'Brien, Race against time: The inside story of Ireland's COVID tracker app (2020). https://www.irishtimes.com/business/technology/race-against-time-the-inside-story-of-ireland-s-covid-tracker-app-1.4303509
44. OHDSI. OMOP common data model (2019). https://www.ohdsi.org/data-standardization/the-common-data-model/
45. OWL Working Group et al. Owl 2 web ontology language document overview: W3C recommendation 27 October 2009. *W3C Recommendation* (2009)

46. A. Ryan, Towards semantic interoperability in healthcare: ontology mapping from SNOMED-CT to hl7 version 3, in *Proceedings of the Second Australasian Workshop on Advances in Ontologies-Volume 72*. Citeseer (2006), pp. 69–74
47. R. Shearer, B. Motik, I. Horrocks, Hermit: A highly-efficient owl reasoner, in *Owled*, vol. 432 (2008), p. 91
48. B. Smith, C. Welty, Ontology: Towards a new synthesis, in *Formal Ontology in Information Systems*, vol. 10(3), pp. 3–9 (ACM Press, New York, 2001)
49. B. Spring, Nursing 2.0: Care in the age of technology (2020). https://ce.ohnurses.org/products/nursing-20-care-in-the-age-of-technology
50. H. Stenzhorn, S. Schulz, M. Boeker, B. Smith, Adapting clinical ontologies in real-world environments. J. Univ. Comput. Sci. **14**(22), 3767 (2008)
51. The guardian. Abandoned NHS it system has cost £10bn so far (2013). https://www.theguardian.com/society/2013/sep/18/nhs-records-system-10bn
52. L.J. Trotter, *More Than Medicine: Nurse Practitioners and the Problems They Solve for Patients, Health Care Organizations, and the State* (Cornell University Press, Ithaca, 2020)
53. United Nations. Sustainable development goals knowledge platform. *United Nations Report* (2015)
54. World Health Organization et al. Who framework on integrated people-centred health services. Technical report, World Health Organization, 2016
55. World Health Organization et al. Shanghai declaration on promoting health in the 2030 agenda for sustainable development. Health Promot. Int. **32**(1), 7 (2017)
56. World Health Organization et al. A coordinated global research roadmap. Technical report, World Health Organization, 2019
57. World Health Organization et al. State of the world's nursing report - 2020. Technical report, World Health Organization, 2020

Chapter 2
Big Data Analytics for Healthcare Information System: Field Study in an US Hospital

Liuliu Fu, Wenlu Zhang, and Lusi Li

2.1 Introduction: IT in US General Hospitals

In the USA, a hospital is often associated with a medical group and it is run by a set of general practitioners, including doctors, nurses, and laboratory technicians. Simultaneously, it has also been widely recognized that Information Technology (IT) market is growing dramatically in recent few years. Combining this, the key role that information plays in health care cannot be ignored. IT costs on healthcare have become a foremost concern of the US government. Health Information Technology (HIT) or Health Information System (HIS) is defined as the computer applications for the practice of medicine [81]. HIS/HIT covers a wide range of applications, such as the Electronic Medical Record (EMR), the Electronic Health Record (EHR), Continuity of Care Document (CCD), Computerized Physician Order Entry (CPOE), decision support systems to assist clinical decision making, and computerized entry systems to collect and storage patient data. According to the report of the U.S. Congressional Budget Office, the Bush Administration established the position of National Coordinator for HIT in the Department of Health and Human Services in 2004 and set the goal of making EHR available to most Americans by 2014. The time to achieving the goal has been revised [15]: in 2008, less than 10% of US hospitals had adopted Basic EHR system; and however, this increased to 76% in 2014. Almost all hospitals (97%) have adopted a certified

L. Fu (✉) · L. Li
California State University, Los Angeles, CA, USA
e-mail: lfu8@calstatela.edu; lli57@calstatela.edu

W. Zhang
California State University, Long Beach, CA, USA
e-mail: Wenlu.Zhang@csulb.edu

S. Paul et al. (eds.), *Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-77558-2_2

EHR technology in 2015, increasing by 35% comparing with 2011. Current data suggests that HIS/HIT has gained increasing recognition in the USA and it is playing a more and more important role for US hospitals [35].

Not only the US government, many leading business companies also realize the potential of HIS/HIT development. Google Health, introduced by Google in 2008 and cancelled in 2011, was a personal health information centralization service that allowed patients to import personal medical records, schedule appointments, and refill prescriptions [96]. As the most similar competitor of Google Health, HealthVault, developed by Microsoft, is a web-based platform where users can see, use, add, and interact with other personal devices such as Windows, Windows phone, iPhone [69]. Microsoft HealthVault allows individuals to manage personal health data via health apps and personal health devices. Intel is now making efforts on multiple perspectives to promote the development of HIS/HIT, including personalized medicine, mobility, devices and imaging, privacy and security, secure cloud [38]. IBM's Healthcare solution aims to enable advanced business models to reduce costs, to create new forms of cooperation, and to promote engagement among business and individuals to increase healthcare outcomes [37]. Subsequently, HIS/HIT has gained visible achievements and is still evolving.

Government and business company efforts bring huge investments into healthcare information systems research in the USA and all over the world. Despite the enormous cost to the hospitals, the overall benefits and costs of HIS have not been deeply assessed [27]. In recent years, much research efforts investigated the link between the implementation of information systems and the performance of organizations. Because hospitals are at the frontier of technology adoption, IT investment becomes one of the main costs of its spending [84]. Many previous studies have indicated a positive relationship between the use of IT and hospital performance [3, 21, 61], but the mechanisms by which IT impacts hospital performance are still not clear: Do HIS/HIT systems influence different hospitals the same way? How to understand and explain the mechanism that HIS/HIT improves the performance of hospitals?

2.2 Overview of Current Healthcare Information Systems

A healthcare system, sometimes referred as "health care system" or "health system," is the integration of people, institutions, and resources that provide health care services. According to the World Health Organization (WHO)'s definition [80]:

> A health system consists of all organizations, people and actions whose primary intent is to promote, restore or maintain health. This includes efforts to influence determinants of health as well as more direct health-improving activities. A health system is therefore more than the pyramid of publicly owned facilities that deliver personal health services. It includes, for example, a mother caring for a sick child at home; private providers; behavior change programmers; vector-control campaigns; health insurance organizations; occupational health and safety legislation. It includes inter-sectoral action by health staff,

for example, encouraging the ministry of education to promote female education, a well-known determinant of better health.

The WHO's definition highlights the fact that there are not only factors of technology, but also factors of human and organization in a healthcare system. All these factors simultaneously determine the outcome of a health care system. In our research, we narrow the broad scope of "system" and define healthcare information systems as computerized systems that facilitate the information sharing and processing within healthcare facilities. Healthcare information systems are fundamentally different from industrial and consumer products which are concerned about market share protection [65]. They need to be able to be implemented across the platforms, and thus there is a requirement for standardization. In general, it has special needs in terms of security, database design, and standards issue.

Evaluating, designing, and implementing HIS/HIT systems cover a wide scope. The key is to integrate the technology factors (e.g., information integration and knowledge management) and social factors (e.g., management, psychology and policy). This multidisciplinary research has drawn interests from many fields including those working in the fields of information system, computer science, business management, medical science, and others. For example: Wilton and McCoy [108] introduced a distributed database which established data links between different applications running in a local network [108]. Both patient information and reference materials were included in their database. Lamoreaux [59] described a database architecture in a medical center in Virginia which integrated the patient treatment file, outpatient clinic file, and fee basis file all together [59]. Johnson et al. [43] discussed the generic database design for patient management information [43] and indicated that the database design needed to allow efficient access to clinical management events from patient, even, location, and provider. Tsumoto [101] developed a rule instruction system to automatically discover the knowledge from an outpatient healthcare system [101], similar to Khoo et al. [52]'s knowledge extraction and discovery system while using the graphical pattern of a medical database [52]. Chandrashekar et al. [14] talked about the considerations when designing a reusable medical database, including the contract issue between the clinical applications and the storage component, multimodality support, centralizing external dependencies, communication models, and performance considerations [14]. Xu et al. [111] introduced an integrated medical supply information system which integrated the demand, service provided, health care service provider's information, inventory storage data, and support tools all together [111]. A recent study by Honglin et al. proposed multiple factor integration (MFI) method to calculate the similarity map for sentence aligning for medical database [110].

With the emergence of these advanced HIS/HIT systems, some well-developed ones have gained wide adoption. Electronic Medical Record (EMR), Electronic Health Record (EHR), and Electronic Patient Record (EPR) are three of the main types adopted. All three systems aim to represent the data electronically and are often used interchangeably. However, fundamental differences exist among these three systems. EMR is the electronic medical information file that is generated during the process of diagnosis. EMR is normally designed according to the diagnosis process in a medical facility, and it is rarely extended outside the scope of a hospital, clinic or medical center. On the other hand, EHR is the systematic collection of electronic health information about patients, which can go beyond the scope of a single medical facility. Thus, EHR integrates information across different facilities and systems, and EMR can serve as a type of data source for the EHR [33, 53]. The scope and purpose of EHR are given by [1]: "a repository of information regarding the health status of a subject of care in computer processable form, stored and transmitted securely, and accessible by multiple authorized users. It has a standardized or commonly agreed logical information model which is independent of EHR systems. Its primary purpose is the support of continuing, efficient and quality integrated health care and it contains information which is retrospective, concurrent and prospective." And finally, EPR refers to "An electronic record of periodic health care of a single individual, provided mainly by one institution" [24], as defined by National Health Service (NHS). The definition of EPR is patient centric. It is the health record of a person along his/her life. NHS has classified EPR into six levels. The research of HIS/HIT may focus on any of the six levels. Our field study looked at Level 1 and 2 of these six levels, which provide evidence to support the intergraded systems on upper levels.

Level 1 - Patient Administration System and Departmental Systems
Level 2 - Integrated Patient Administration and Departmental Systems
Level 3 - Clinical activity support and noting
Level 4 - Clinical knowledge, decision support, and integrated care pathways
Level 5 - Advanced clinical documentation and integration
Level 6 - Full multimedia EPR on line

From the perspective of information location, content, source, maintainer, and user, we compare EMR, EHR, and EPR in Table 2.1:

Although these well-developed systems have gained wide acceptance and have been implemented by most healthcare facilities today, many studies have discussed the issues regarding the implementation of the EMR/EHR/EPR as well as the problems of the system design. For example, some studies discussed the accuracy issues of quantitative EMR data [17, 30, 98, 104]. Particularly, Wagner and Hogan indicated that the main cause of errors was the failure to capture the patient's mistake when misreporting about medications, and the second most important cause for the error was the failure to capture medication changes from outside clinicians. Moody et al. [72] found that only small amount of nurses reported that EHRs had resulted in a decreased workload, while the majority of nurses preferred bedside documentation [72]. Bygholm [12] found the implementation issues of EPR systems from a case study [12], and it was argued that there was a need to distinguish different types of end-user support when various types of activity were involved.

Table 2.1 Comparison among EMR, EHR, and EPR

	EMR	EHR	EPR
Purpose	Managerial process control on a medical domain	Information sharing	Personal health management
Information allocation	Health facilities	Public health department	Individual person
Information content	Medical record	Medical record and public health record	Medical record and personal health record
Information control	Health practitioner or related stuff can gain access	Health practitioner, related stuff in the health facilities, and government stuff can gain access	Can get access only after get permission by the record owner
Information resource	Single health facilities	Multiple health facilities	Single Health facility and individuals
Information maintainer	Health facility	Government	Individual

2.3 The Measurement of the Healthcare System

Performance measurement is defined as "the process of quantifying the efficiency and effectiveness of action," or "a metric used to quantify the efficiency and/or effectiveness of an action," or "the set of metrics used to quantify both the efficiency and effectiveness of actions" [76], [78]. Here three main issues are covered: "quantification," "efficiency and effectiveness," and "metrics." Quantification means that the results of performance measurement need to be countable and comparable. Efficiency and effectiveness are the measuring objects. Metrics emphasize that performance measurement is multidimensional.

In most cases, the process of measuring performance requires the uses of statistical tools to determine results. Today many performance measurement systems have gained great achievements. For example, the Balanced Scorecard, first proposed in 1992, provides a comprehensive framework to translate a company's strategic objectives into a related set of performance measures [49, 50], including the financial perspective, customer perspective, internal business perspective, and innovation and learning perspective. Neely's "Performance Prism" system looks at five interrelated facets of the prism: stakeholder satisfaction, stakeholder contribution, strategies, process, and capabilities [2, 77, 79]. More detailed measuring perspectives are defined under each facet. The Performance Pyramid developed by Lynch and Cross contains a hierarchy of financial and nonfinancial performance measures. The four-level pyramid system shows the link between strategies and operations, translating the strategic objectives top down, and rolling measures bottom up [18]. Dixon [22] developed the Performance Measurement Questionnaire (PMQ) system to determine the degree that the existing performance measures supported the improvements and to identify what the organization needed for improvement

[22]. For team-based structures, Jones and Schilling [44] proposed the approaches of the Total Productive Maintenance (TPM) process in which a practical guide for developing a team's vital measurement system is provided [44]. Later after the proposition of TPM, the 7-step TPM process [62] and Total Measurement Development Method (TMDM) [31] were developed. By studying the processes and strategies with organizations, these systems function as a part of the management process giving insights on what should be achieved and whether the outputs meet intended goals.

Since performance measurement is multidimensional, a Performance Measurement System (PMS) can differ when the situation and context change. Despite the variety of PMSs, some universal steps and requirements need to be followed when designing a meaningful measurement system. Three general steps are included when designing a performance measurement system: defining strategic objectives, deciding what to measure, and installing performance measurement system into management thinking [51]. Wisner and Fawcett later added more operational details into the procedure, expanding the three steps to a nine-step flow diagram [109].

Particularly for the measurement of healthcare related systems, Purbey et al. adopted Beamon's evaluation criteria for supply chain performance [8], coming up with a set of measurement characteristics for healthcare processes: inclusiveness, universality, measurability, consistency, and applicability [87]. Due to the complexity of healthcare systems, there are various aspects implicating the system performance. Looking at the review of Van Peursem et al., three measurement groups are included for health management performance: (1) economy, efficiency, and effectiveness; (2) quality of care; and (3) process [102]. These measurement aspects focused on the quality of management, not the quality of medical practice. The first aspect mentioned here (economy, efficiency, and effectiveness) is normally referred to as the three e's and it has been devised for public sector organizations [11, 66, 70]. A PMS for HIS/HIT can also be classified as financial or nonfinancial [68, 90, 102]. Table 2.2 summarizes the studies on healthcare system performance and their measurements according to financial and nonfinancial categories:

As a short conclusion to this section, existing healthcare systems have gained long-term success, while there remain many unsolved issues regarding the implementation and use of such systems. More research needs to be done to improve the usability and data quality of healthcare systems. There is a high demand for a further investigation of current system's weaknesses and the development of integrated healthcare systems. As a result, we conduct this field study to collect evidence from a general hospital. The information we gathered contains both the qualitative and quantitative ones. Before looking at the details of the field study, let's compare these two general research categories: quantitative research and qualitative one.

Table 2.2 Healthcare system studies with financial and nonfinancial measurements

Financial measurement	Nonfinancial measurement
Return on investment (ROI) [67] Medicaid inpatient revenue [29] Total income/revenue [4] Cost, market share grow, return on assets (ROA), ROI, operating profit [63] ROA, operating margin, market share, sales growth, current ratio, debt ratio, cash flow to debt ratio, cumulative depreciation ratio [41] Net operating revenue, market share, total margin, total revenue [58] ROA, operating margin, net cash flow, adjusted net patient revenue (Bill Binglong [106])	Patient satisfaction [10, 13, 85, 86] Patient safety (Bill Binglong [106]) For three clinical areas: hip/knee surgery, cardiac care, and obstetric care, hospitals were rated as better than expected (fewer deaths/ complications), as expected, or worse than expected [36] Standardized mortality ratio (SMR) [39, 45, 71, 91] Bed Occupancy Rate (BOR) [4] Mortality, readmission, and complication [20] Percent occupancy [58]

2.4 Quantitative Research

Quantitative research methods are rooted in the natural sciences [74]. The objective is to measure a particular phenomenon using quantified datasets of a chosen sample from the population of interest. In general, using quantitative methods requires the inclusion of a large sample size in order to fully represent the population of interest. Sometimes quantitative research can be followed by qualitative research to further investigate the details of some findings, or it can follow qualitative research in order to prove the validity of proposed assumptions. Quantitative research methods are widely accepted in the field of social science. There are several examples of application of quantitative methods in HIS/HIT studies.

Mathematical modeling [9, 56, 114] means to construct and describe a system using mathematical concepts and equations.

Experimental method in information system studies is a controlled procedure in which independent variables are manipulated by the researchers, and the dependent variable is measured to test the hypotheses [26, 28, 54].

Survey method [7, 89, 93, 105] studies the sampling of datasets from a population using collected survey data. A survey can be cross-sectional (collecting data from people for one time) or longitudinal (collecting information from the same people over time). The cross-sectional method simply measures the research subjects without manipulating the external environment. If multiple groups are selected, it can compare different population groups at a single point of time. In contrast, longitudinal survey method collects information from multiple time frames. It has a significant advantage over cross-section methods in identifying cause-and-effect relationships. However, longitudinal survey method also faces the challenges associated with following a study group over a long time period.

Quantitative methods are most suitable when a researcher wants to know “how much”: the size and extent or duration of certain phenomena [94]. Especially when testing the cost, quality, or performance of HIS/HIT systems, quantitative methods become a main choice of evaluation. For instance, to evaluate the financial performance of HIS/HIT systems, quantitative methods are suitable to use. One of the main strengths of quantitative approaches is their reliability and objectivity. With a well-constructed analytical model, they are able to simplify a complex problem to a limited number of variables. This requires establishing the testing model prior to data collection and the collected data to be precise and able to reflect the target population. Once the data collecting process is complete, data analysis becomes relatively less time-consuming especially with the help of statistical software (e.g., SPSS, Matlab, Minitab, SAS, Excel). What one needs to note is that the research results are relatively independent of the researchers. For example, researchers cannot guarantee whether the outputs are statistically significant, or whether the model fit can be proved. There are also some weaknesses of quantitative methods. As the tested models are constructed before data collection, the researchers might miss some important factors of the phenomena, because the focus is "hypotheses testing" rather than "hypotheses generation" (R. B. [42]). Therefore, the tested model needs to be reasonable and with a valid theoretical background.

2.5 Qualitative Research

In contrast to quantitative ones, qualitative research methods were originally developed for the social sciences [74] who are concerned with "developing explanations of social phenomena [34]." The purpose of utilizing qualitative methods is to gain an in-depth understanding of underlying factors and to uncover hidden trends. More importantly, they are able to provide insights and ideas for future quantitative research: to determine not only what is happening, or what might be important to measure, but why to measure and how people think or feel [48]. Unlike quantitative methods that require large number of datasets in general, qualitative methods usually concentrate on a small number of cases. Examples of qualitative approaches in the field of information systems given by Myers are action research, case study research, and ethnography [74].

Action research "seeks to bring together action and reflection, theory and practice, in participation with others, in the pursuit of practical solutions to issues of pressing concern to people, and more generally the flourishing of individual persons and their communities." [88]. By this definition, action research method for HIS/HIT has its concern on the perspective of human and organizational factors. Reason and Bradbury concluded that action research could be an ideal postpositivist social scientific research method in information system discipline [88].

Case study research methods intend to implement up-close and detailed examination of a subject of the case. They are analyses of person, projects, periods, policies, decisions, events, institutions, or other systems that are under the study by one or more methods (G. [99]). By its nature, the case study approach can be applied on almost all perspectives of HIS/HIT research. Many cases are presented all over the world, such as the United States [47], Australia [23], Netherland [103], Taiwan [107], Philippines [40], and Africa [46].

The word *ethnography* has its origin in Greek where ethnos means "folk, people, nation" and *grapho* means "I write" [95]. The goal of ethnography research is to improve people's understanding of human thought and activities via investigation of human actions in context [73]. Therefore, ethnography approaches in HIS/HIT research also focus on the social aspects of the field, for instance, organizational culture [5], power and managerial issues [75], and to contribute to the design process drawing examples to build explanation system [25].

Unlike quantitative approaches which check comparatively large sample sizes, qualitative approaches examine specific cases. It is useful when investigating complex situations involving a limited number of cases, and it provides rich detail of the phenomena in specific contexts. Quantitative approaches require data standardization in order to process and compare statistical results, while qualitative approaches allow the researchers to explore the responses as they are and to observe the behaviors, opinions, needs, and patterns without yet fully understanding whether the data are meaningful or not [64]. As a result, they are able to help HIS/HIT researchers capture some important hidden factors which might be ignored with quantitative approaches. However, because of the flexibility of the collected data, it takes more time for data processing and data analysis. Moreover, the results of interpretation and quality are easily influenced by researchers' personal knowledge and biases. Therefore, qualitative and quantitative methods have been integrated in many HIS/HIT studies to compensate each other.

2.6 Challenges in Understanding Existing Healthcare Information Systems

Identifying the challenges means to explore the influences from the physical, socioeconomic, and work environments [92]. One of the most widely studied questions regarding the performance of current systems is: what matters? These factors can relate to multiple perspectives such as human, organization, and technology. We find a lot of influential factors under different contexts, for example:

- Staff and clinic size, doctor waiting time, the use of appointment scheduler (new or follow-up patient) [16]
- Time interval until the next appointment, doctor number, keep record of follow-up patient, improve the communications, booking no routine patients for the first 45 minutes for each clinic, field-of-vision appointments before 1st appointment, redesign the appointment card to give patients more information about their next visit to clinic [9]
- Number of operators, registration windows, physicians nurses, medical assistants, check in rooms, specialty rooms [97]
- Appointment scheduling for no-shows. Solution: overbooking [56]

- Appointment scheduling, appointment supply and consumption process, no-shows, overbooking [55]
- Different appointment types, no shows, overbooking [32]
- Length of time patients had attended the clinic, patients' mode of transport to the clinic [100]

Now the challenge is not only whether the factors matter or what factors matter, but also at which level they matter, and why they matter. Lau's review on HIS research summarized the factors of HIS studies into Information System Success Model [19, 60]. It is clear that understanding HIS/HIT systems is multidisciplinary. As discussed earlier, the research scope of HIS/HIT covers the aspects of technology, organization, social, and human. To evaluate the quality or performance of an existing health information system, we need to include elements from multiple perspectives: technical factors (such as information quality, system easiness of use, system reliability and response time), social factors (such as policy enforcement), and financial factors (such as different types of costs). We will start from a field study to explore the environment of healthcare providers and to collect information about doctors and patients.

2.7 The Field Study of EVMS

The Eastern Virginia Medical School (EVMS) clinic is located on South Hampton Avenue, Norfolk, Virginia, USA. It belongs to Denver Community Health Services (DCHS), which is a network of 8 community health Centers, 12 school-based health centers, and 2 urgent care centers. The EVMS research provides coordination for research committees as well as the research advisory group. It wins funding from outside sources, with grants and contracts awarded for work in Hampton Roads communities. EVMS has a long-term cooperation with our research team and provided great support for this field study.

The physicians of EVMS specialize in family and internal medicine, obstetrics, medical and surgical specialties as well as radiation oncology, laboratory, and pathology services, with the mission "to provide patient-centered quality healthcare to the patients that we serve." In order to reach the goal, the medical group has been working very hard to deliver care that is safe, efficient, and cost-effective. By the time of our investigation, EVMS had been using patient portal to keep records. In order to explore the current situation of EVMS Ghent Family Medicine, we conduct a data analysis to identify the discrepancy between patient demand and provider supply, to see whether the capacity management in such an outpatient family machine has brought a good outcome and whether the patient data in the HIT/HIS system was utilized effectively.

The datasets from EVMS were mainly drawn from the scheduling record spreadsheet provided by the hospital. Some data came from our interview with the doctors, such as the general workloads of doctors and residents. The dataset consists of the doctor's schedule and patient records from July 2012 to December 2012. There are 131 days, for both morning and afternoon schedules. In our analysis, we take the average of the doctor and patient number for the morning and afternoon as the data points. Our results are as follows.

2.8 Statistical Findings

Figure 2.1 shows a linear relationship between the number of doctors and the number of the patient. According to Fig. 2.1, each doctor takes care of 6 to 7 patients in 4 hours (half a day) on average.

The actual workload for each doctor is 6 to 7 patients per day, according to our collected data. On the other hand, our interview also indicates that each doctor spends about 20 minutes on a single patient on average. Therefore, the actual total workload is less than what the doctors are able to handle. We question that it might be the poor design of scheduling system that reduces the efficiency. There is room for improvement.

From Table 2.3, we can see that Tuesdays and Wednesdays are easy days, while Mondays, Thursdays, and Fridays are busy days, especially on Mondays. Moreover, the standard deviation associated with patients is much higher than that of doctors every day, especially on Mondays. Then here comes the question: does the current schedule respond to the high demand on Mondays?

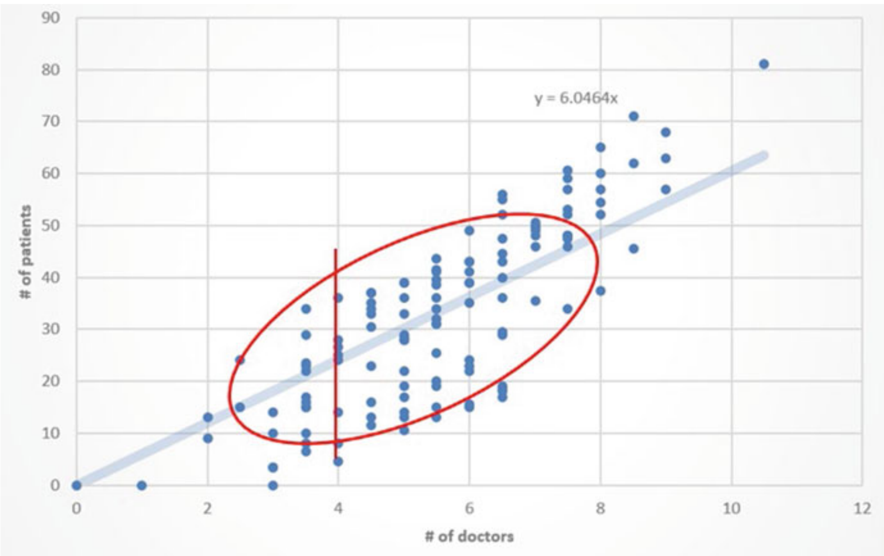


Fig. 2.1 Relationship between number of doctors and number of patients (force $s = 0$)

Table 2.3 Number of patients and doctors each day (half day based, holiday excluded)

	Average # of doctors	STD DEV of doctors	Average # of patients	STD DEV of patients	# of pts. per doc
Monday	7.4	1.3	51.3	13.0	7
Tuesday	6.1	1.3	33	13.0	5.5
Wednesday	4.6	1.4	12.5	5.2	2.7
Thursday	5.1	1.3	33.7	11.8	6.6
Friday	4.8	1.64	32.6	15.3	6.8

Table 2.4 Number of patients and doctors each month (half day based, holiday excluded)

	Average # of doctors	STD DEV of doctors	Average # of patients	STD DEV of patients	# of pts. per doc
July	5.2	1.9	31.1	16.4	6.0
August	5.0	1.1	29.6	13.3	5.9
September	6.6	1.4	38.5	17.0	5.8
October	5.3	1.6	27.5	16.9	5.2
November	6.0	2.4	38.7	21.8	6.5
December	5.1	1.9	29.2	16.7	5.7

A similar pattern is also found when we do monthly demand analysis (Table 2.4): November has the highest standard deviation associated with the patient as well as the doctors. The assumption is that it is because of the seasonal factors: November is the month of Thanksgiving and it is very close to Christmas break. People tend to travel, have parties, reunion, and engage in more risky behavior in terms of health issues. Thus, it has the highest variation in demand.

2.9 Gap Between Patient Demand and Doctor Schedule

Figure 2.2 shows the changes in the patient numbers and doctor numbers in half a year. We can see that the service time provided by physicians is level and stable, while the demand for service from patients is sporadic and lumpy. Figure 2.2 suggests that sometimes there were too many service hours, and at other times there appeared to be insufficient service resources that might lead to long waiting times and unhappy patients. Delays in obtaining service lead to patient dissatisfaction, higher cost, and adverse consequences. Similarly, comparing with the actual number of patients seen by the doctors each day which is sporadic and lumpy in Fig. 2.3, the line for the expected number of patients appears more level and stable. It indicates that the current patient schedule does not fit the intended workload capability of doctors.

Finally, we face such a question: are we able to determine a consistent demand pattern that matches the level supply of providers? What we find is that the pattern of

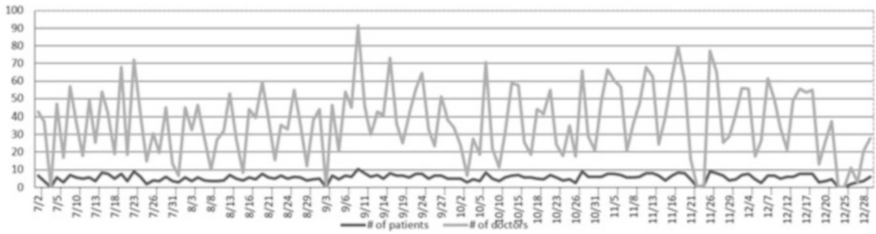


Fig. 2.2 Number of patients and doctors of the time period

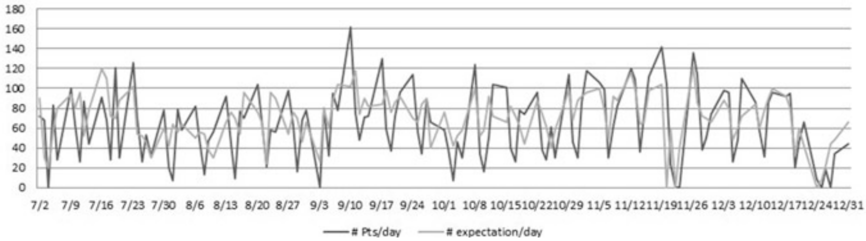


Fig. 2.3 The actual number of patients seen each day versus the expected number of patients seen each day

the patient demand and the service provider is not consistent. As shown in Fig. 2.4, the shape of the demand and service curve can be a triangle, a negative slope, and a concave. Other than these standard shapes, there are some other shapes as shown in Fig. 2.4(d). In other words, the variability of patient demand and the service seems to be significant.

Such variability may come from patients and service providers. From the perspective of patients, the variability comes from (1) different patient types, such as new patients, follow-up patients, return patients; (2) different schedule types, such as by appointments, late show, no show, overbooking, walk-in patients, urgent patients, emergencies, patients who want the same doctor; and (3) different service times, such as the diagnosis by annual physical, for new patients, for follow-ups, for patients who want to have all health issues done in one visit. From the perspective of the service providers, the variability may come from (1) the difference in provider's schedule, for example, the doctor schedule is made quarterly, 3 ~ 4 months in advance, while the medical aid schedule is made a day before the service; (2) variability in-service time, that the standard (20 minutes per patient) does not apply to all doctors and there is at least a 5% chance the doctors will run their appointment late. Our findings highlight the mismatch between patient demand and the schedule of the service provider.

Our goal is to reduce the bottleneck of the services, reduce the waiting time of the patients, and improve patients' satisfaction towards the services. Some lean service operations can take place to reach the goal, such as better scheduling, understanding

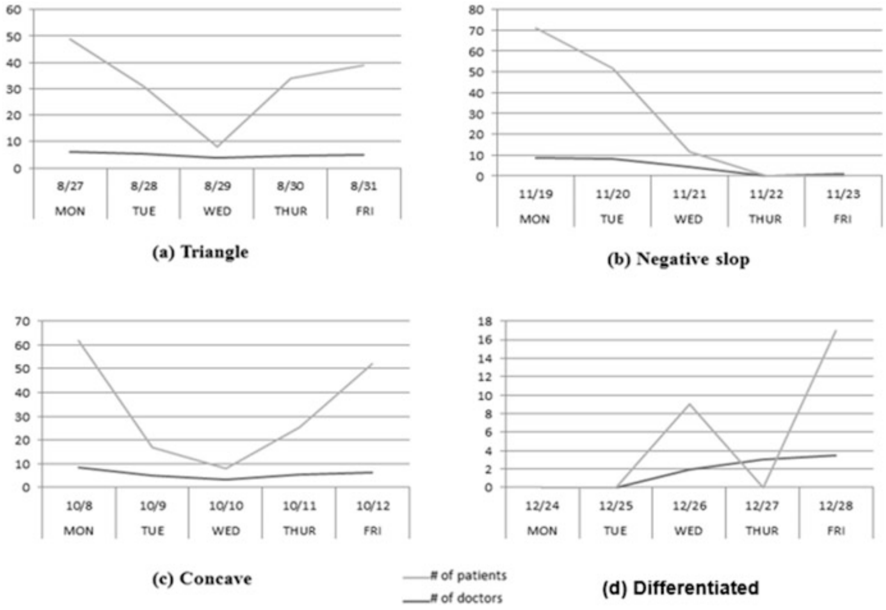


Fig. 2.4 The pattern of patient demand and service provider by weeks

the patient's needs and their tolerance span, and matching patient's demand with providers' supply. For example, parents with young children will be scheduled early in the morning or late in the afternoon, so the parents do not need to take time off during the day; retired senior citizens (who do not mind waiting a little longer than the scheduled time) can be scheduled in the middle of the day. The physician schedule, nurse scheduling, and patient schedule need to be integrated, and the patient information also needs to be integrated with the staff schedule. Such categorizing functions can be performed by decision support systems. It is one of our following studies that fall into the categories of level 3 and 4 in NHS framework.

2.10 Discussion and Future Work

The field study in EVMS highlights the mismatch between patient demand and service provider schedules. This was actually the very 1st step of our series of study. One of the following works was to adopt the institutional theory to explain the process of implementing HIS/HIT and the possible outcomes. Using HIMSS data, we run structural equation model to test six hypotheses to identify the relationships between size and service volume, size and performance, and size and IT implementation. To solve the problem of existing systems, we will conduct further research to adopt decision support methods to capture the classification patterns from the doctor. Such a system should be able to provide valuable recommendations to health providers, helping them gain more transparent information from patients, and make better scheduling decisions to minimize gaps between patient demand and the provided services.

This field study evaluated Level 1 and 2 of NHS model. It reveals that the physicians in the EVMS hospital were not satisfied to the current scheduling status by the time the survey was conducted. The studies with SERVQUAL, a measurement framework for service quality [83], also show similar findings. SERVQUAL that evaluates the quality of clinic services based on information systems will be measured and compared from five dimensions: tangibles, reliability, responsiveness, assurance, and empathy:

- Tangibles: physical facilities, equipment, and appearance of personnel
- Reliability: ability to perform the promised service reliable and accurately
- Responsiveness: willingness to help customers and provide prompt service
- Assurance: knowledge and courtesy of employees and their ability to inspire trust and confidence
- Empathy: caring, individualized attention provided to customers

Many studies have gained success in adopting the SERVQAUL model to evaluate the performance in health care research discipline. Babakus and Mangold [6] found that the SERVQUAL scales could be used to assess the gap between the patient perceptions and expectations and that SERVQUAL was applicable as a standardized measurement scale to compare results in different industries [6]. In particular, Lam [57] checked a hospital service quality in Hong Kong and the result indicated that SERVQUAL was consistent and reliable as a measurement tool [57]. Youssef et al. [113] examined at the service quality of NHS hospitals [113]. Pakdil and Harwood evaluated the patient satisfaction for a preoperative assessment clinic with SERVQUAL [82]. And a recent study in 2010 compared the service quality between public and private hospitals using SERVQUAL [112]. Based on all these facts, we can also adopt SERVQUAL as a reliable measurement infrastructure for our proposed healthcare systems of following studies.

References

1. ISO/TR 20514. Health Informatics-Electronic Health Record-Definition, Scope, and Content. In: ISO Geneva (2005)
2. C. Adams, A. Neely, The performance prism to boost M&A success. Meas. Bus. Excell. **4**(3), 19–23 (2000)
3. H. Ahmadi, M. Nilashi, L. Shahmoradi, O. Ibrahim, F. Sadoughi, M. Alizadeh, A. Alizadeh, The moderating effect of hospital size on inter and intra-organizational factors of hospital information system adoption. Technol. Forecast. Soc. Chang. **134**, 124–149 (2018)
4. H. Akashi, T. Yamada, E. Huot, K. Kanal, T. Sugimoto, User fees at a public hospital in Cambodia: Effects on hospital performance and provider attitudes. Soc. Sci. Med. **58**(3), 553–564 (2004)

5. D.E. Avison, M.D. Myers, Information systems and anthropology: And anthropological perspective on IT and organizational culture. Inf. Technol. People **8**(3), 43–56 (1995)
6. E. Babakus, W.G. Mangold, Adapting the SERVQUAL scale to hospital services: An empirical investigation. Health Serv. Res. **26**(6), 767 (1992)
7. L. Baker, T.H. Wagner, S. Singer, M.K. Bundorf, Use of the internet and e-mail for health care information: Results from a national survey. JAMA **289**(18), 2400–2406 (2003)
8. B.M. Beamon, Measuring supply chain performance. Int. J. Oper. Prod. Manag. **19**(3), 275–292 (1999)
9. J.C. Bennett, D. Worthington, An example of a good but partially successful OR engagement: Improving outpatient clinic operations. Interfaces **28**(5), 56–69 (1998)
10. W. Boulding, S.W. Glickman, M.P. Manary, K.A. Schulman, R. Staelin, Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days. Am. J. Manag. Care **17**(1), 41–48 (2011)
11. S. Brignall, S. Modell, An institutional perspective on performance measurement and management in the 'new public sector'. Manag. Account. Res. **11**(3), 281–306 (2000)
12. A. Bygholm, End-user support: A necessary issue in the implementation and use of EPR systems. Stud. Health Technol. Inform. **84**(Pt 1), 604–608 (2000)
13. R.A. Carr-Hill, The measurement of patient satisfaction. J. Public Health **14**(3), 236–249 (1992)
14. N. Chandrashekar, S. Gautam, K. Srinivas, J. Vijayananda. *Design considerations for a reusable medical database.* Paper presented at the Computer-Based Medical Systems, 2006. CBMS 2006. 19th IEEE International Symposium on (2006)
15. D. Charles, M. Gabriel, T. Searcy, *Adoption of electronic health record systems among US non-federal acute care hospitals: 2008–2014* (Office of the National Coordinator for Health Information Technology, 2015)
16. J.E. Clague, P.G. Reed, J. Barlow, R. Rada, M. Clarke, R.H. Edwards, Improving outpatient clinic efficiency using computer simulation. Int. J. Health Care Qual. Assur. **10**(5), 197–201 (1997)
17. K. Corson, M.S. Gerrity, S.K. Dobscha, J.B. Perlin, R.M. Kolodner, R.H. Roswell, Assessing the accuracy of computerized medication histories. Am J Manag Care ***10***(part 2), 872–877 (2004)
18. K.F. Cross, R.L. Lynch, The "SMART" way to define and sustain success. National Productiv. Rev. **8**(1), 23–33 (1988)
19. W.H. Delone, E.R. Mclean, Measuring e-commerce success: Applying the DeLone & McLean information systems success model. Int. J. Electron. Commer. **9**(1), 31–47 (2004)
20. S.I. DesHarnais, L.F. McMahon Jr., R.T. Wroblewski, A.J. Hogan, Measuring hospital performance: The development and validation of risk-adjusted indexes of mortality, readmissions, and complications. Med Care **28**, 1127–1141 (1990)
21. S. Devaraj, R. Kohli, Performance impacts of information technology: Is actual usage the missing link? Manag. Sci. **49**(3), 273–289 (2003)
22. Dixon, J. R. (1990). *The new performance challenge: Measuring operations for world-class competition*: Irwin Professional Pub
23. M. Evered, S. Bögeholz *A case study in access control requirements for a health information system.* Paper presented at the Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation-Volume 32 (2004).
24. N. Executive. *Information for health: an information strategy for the modern NHS 1998–2005: a national strategy for local implementation*: NHS Executive (1998)
25. D.E. Forsythe, Using ethnography in the design of an explanation system. Expert Syst. Appl. **8**(4), 403–417 (1995)
26. C.R. Franz, D. Robey, R.R. Koeblitz, User response to an online information system: A field experiment. MIS Q., 29–42 (1986)
27. C.P. Friedman, J.C. Wyatt, J. Faughnan, Evaluation methods in medical informatics. BMJ **315**(7109), 689 (1997)

28. L. Fu, K. Maly, E. Rasnick, H. Wu, M. Zubair, User experiments of a social. Faceted Multimedia Classification System
29. G.O. Ginn, R.P. Lee, Community orientation, strategic flexibility, and financial performance in hospitals. J. Healthcare Manage. Am. Coll. Healthcare Execut. **51**(2), 111–121 (2005) discussion 121-112
30. Goldberg, S. I., Shubina, M., Niemierko, A., & Turchin, A. (2010). A Weighty Problem: Identification, Characteristics and Risk Factors for Errors in EMR Data. Paper presented at the AMIA Annual Symposium Proceedings
31. C.F. Gomes, M.M. Yasin, J.V. Lisboa, Performance measurement practices in manufacturing firms: An empirical investigation. J. Manuf. Technol. Manag. **17**(2), 144–167 (2006)
32. M. Guo, M. Wagner, & C. West. *Outpatient clinic scheduling: a simulation approach.* Paper presented at the Proceedings of the 36th conference on Winter simulation (2004)
33. J.L. Habib, EHRs, meaningful use, and a model EMR. Drug Benefit. Trends **22**(4), 99–101 (2010)
34. Hancock, B., Ockleford, E., & Windridge, K. (1998). *An introduction to qualitative research*: Trent Focus Group Nottingham
35. J. Henry, Y. Pylypchuk, T. Searcy, V. Patel, Adoption of electronic health record systems among US non-federal acute care hospitals: 2008–2015. ONC Data Brief **35**, 1–9 (2016)
36. J.H. Hibbard, J. Stockard, M. Tusler, Hospital performance reports: Impact on quality, market share, and reputation. Health Aff. **24**(4), 1150–1160 (2005)
37. IBM. Healthcare (2015). Retrieved from http://www-935.ibm.com/industries/healthcare/
38. Intel. Intel Health and Life Sicences (2015). Retrieved from http://www.intel.com/content/www/us/en/healthcare-it/healthcare-overview.html
39. B. Jarman, D. Pieter, A. van der Veen, R. Kool, P. Aylin, A. Bottle, et al., The hospital standardised mortality ratio: A powerful tool for Dutch hospitals to assess their quality of care? Qual. Safety Health Care **19**(1), 9–13 (2010)
40. R. Jayasuriya, Managing information systems for health services in a developing country: A case study using a contextualist framework. Int. J. Inf. Manag. **19**(5), 335–349 (1999)
41. M. Je'McCracken, T.F. McIlwain, M.D. Fottler, Measuring organizational performance in the hospital industry: An exploratory comparison of objective and subjective methods. Health Serv. Manag. Res. **14**(4), 211–219 (2001)
42. R.B. Johnson, A.J. Onwuegbuzie, Mixed methods research: A research paradigm whose time has come. Educ. Res. **33**(7), 14–26 (2004)
43. Johnson, S. B., Paul, T., & Khenina, A. (1997). Generic Database Design for Patient Management Information. Paper presented at the Proceedings of the AMIA Annual Fall Symposium
44. S. D. Jones, D. J. Schilling *Measuring team performance: a step-by-step, customizable approach for managers, facilitators, and team leaders* (2000). Retrieved from
45. J.M. Kahn, A.A. Kramer, G.D. Rubenfeld, Transferring critically ill patients out of hospital improves the standardized mortality ratio: A simulation study. Chest J. **131**(1), 68–75 (2007)
46. R.M. Kamadjeu, E.M. Tapang, R.N. Moluh, Designing and implementing an electronic health record system in primary care practice in sub-Saharan Africa: A case study from Cameroon. Inform. Prim. Care **13**(3), 179–186 (2005)
47. B. Kaplan, D. Duchon, Combining qualitative and quantitative methods in information systems research: A case study. MIS Q. **12**, 571–586 (1988)
48. B. Kaplan, J.A. Maxwell, Qualitative research methods for evaluating computer information system, in *Evaluating the organizational impact of healthcare information systems*, (Springer, New York, 2005), pp. 30–55
49. R. S. Kaplan, D. P. Norton. Putting the balanced scorecard to work. *Performance measurement, management, and appraisal sourcebook, 66* (1995)
50. R.S. Kaplan, D.P. Norton, The balanced scorecard: Measures that drive performance. Harv. Bus. Rev. **83**(7), 172–180 (2005)
51. D.P. Keegan, R.G. Eiler, C.R. Jones, Are your performance measures obsolete. Manag. Account. **70**(12), 45–50 (1989)

52. C. S. Khoo, S. Chan, Y. Niu. *Extracting causal knowledge from a medical database using graphical patterns.* Paper presented at the proceedings of the 38th annual meeting on Association for Computational Linguistics (2000)
53. P. Kierkegaard, Electronic health record: Wiring Europe's healthcare. Comp. Law Secur. Rev. **27**(5), 503–515 (2011)
54. M. Korpela, H.A. Soriyan, K. Olufokunbi, A. Onayade, A. Davies-Adetugbo, D. Adesanmi, Community participation in health informatics in Africa: An experiment in tripartite partnership in Ile-Ife, Nigeria. Comp. Support. Cooperat. Work (CSCW) **7**(3–4), 339–358 (1998)
55. L.R. LaGanga, Lean service operations: Reflections and new directions for capacity expansion in outpatient clinics. J. Oper. Manag. **29**(5), 422–433 (2011)
56. L.R. LaGanga, S.R. Lawrence, Clinic overbooking to improve patient access and increase provider productivity*. Decis. Sci. **38**(2), 251–276 (2007)
57. S.S. Lam, SERVQUAL: A tool for measuring patients' opinions of hospital service quality in Hong Kong. Total Qual. Manag. **8**(4), 145–152 (1997)
58. B.T. Lamont, D. Marlin, J.J. Hoffman, Porter's generic strategies, discontinuous environments, and performance: A longitudinal study of changing strategies in the hospital industry. Health Serv. Res. **28**(5), 623 (1993)
59. J. Lamoreaux, The organizational structure for medical information management in the Department of Veterans Affairs: An overview of major health care databases. Med. Care **34**(3), 31–44 (1996)
60. F. Lau, C. Kuziemsky, M. Price, J. Gardner, A review on systematic reviews of health information system studies. J. Am. Med. Inform. Assoc. **17**(6), 637–645 (2010)
61. K. Lee, T.T. Wan, Effects of hospitals' structural clinical integration on efficiency and patient outcome. Health Serv. Manag. Res. **15**(4), 234–244 (2002)
62. Leflar, J. (2001). *Practical TPM: successful equipment management at Agilent Technologies*: Productivity Press
63. L. Li, W. Benton, G.K. Leong, The impact of strategic operations management decisions on community hospital performance. J. Oper. Manag. **20**(4), 389–408 (2002)
64. Madrigal, D., & McClain, B. (2012). Strengths and Weaknesses of Quantitative and Qualitative Research. *UX Matters*
65. K.D. Mandl, I.S. Kohane, Escaping the EHR trap—The future of health IT. N. Engl. J. Med. **366**(24), 2240–2242 (2012)
66. D.J. Mayston, Non-profit performance indicators in the public sector. Financial Account. Manage. **1**(1), 51–74 (1985)
67. N. Menachemi, J. Burkhardt, R. Shewchuk, D. Burke, R.G. Brooks, Hospital information technology and positive financial performance: A different approach to finding an ROI. J. Healthcare Manage. Am. Coll. Healthcare Execut. **51**(1), 40–58 (2005) discussion 58-49
68. P. Micheli, M. Kennerley, Performance measurement frameworks in public and non-profit sectors. Product. Plan. Control **16**(2), 125–134 (2005)
69. Microsoft (2015). Retrieved from https://www.healthvault.com/us/en
70. A. Midwinter, Developing performance indicators for local government: The Scottish experience. Public Money Manage. **14**(2), 37–43 (1994)
71. A.J. Molyneux, R.S. Kerr, J. Birks, N. Ramzi, J. Yarnold, M. Sneade, et al., Risk of recurrent subarachnoid haemorrhage, death, or dependence and standardised mortality ratios after clipping or coiling of an intracranial aneurysm in the International Subarachnoid Aneurysm Trial (ISAT): long-term follow-up. Lancet Neurol. **8**(5), 427–433 (2009)
72. L.E. Moody, E. Slocumb, B. Berg, D. Jackson, Electronic health records documentation in nursing: nurses' perceptions, attitudes, and preferences. Comp. Inform. Nursing **22**(6), 337–344 (2004)
73. M.D. Myers, Critical ethnography in information systems, in *Information systems and qualitative research*, (Springer, Boston, 1997a), pp. 276–300
74. M.D. Myers, Qualitative research in information systems. Manag. Inf. Syst. Q. **21**, 241–242 (1997b)

75. M.D. Myers, L.W. Young, Hidden agendas, power and managerial assumptions in information systems development: An ethnographic study. Inf. Technol. People **10**(3), 224–240 (1997)
76. Neely, A. (1994). Performance measurement system design–third phase. Performance Measurement System Design Workbook*,* 1
77. A. Neely, C. Adams, P. Crowe, The performance prism in practice. Meas. Bus. Excell. **5**(2), 6–13 (2001)
78. A. Neely, M. Gregory, K. Platts, Performance measurement system design: A literature review and research agenda. Int. J. Oper. Prod. Manag. **15**(4), 80–116 (1995)
79. Neely, A. D., Adams, C., & Kennerley, M. (2002). *The performance prism: The scorecard for measuring and managing business success*: Prentice Hall Financial Times London
80. Organization, W. H. (2007). Everybody's Business–Strengthening Health Systems to Improve Health Outcomes: WHO's Framework for Action
81. Orszag, P. R. (2008). Evidence on the Costs and Benefits of Health Information Technology. Paper presented at the Testimony before Congress
82. F. Pakdil, T.N. Harwood, Patient satisfaction in a preoperative assessment clinic: An analysis using SERVQUAL dimensions. Total Qual. Manag. Bus. Excell. **16**(1), 15–30 (2005)
83. A. Parasuraman, V.A. Zeithaml, L.L. Berry, Servqual. J. Retail. **64**(1), 12–37 (1988)
84. S.T. Parente, R.L. Van Horn, Valuing hospital investment in information technology: Does governance make a difference? Health Care Financ. Rev. **28**(2), 31–43 (2005)
85. G.C. Pascoe, Patient satisfaction in primary health care: A literature review and analysis. Eval. Program Plann. **6**(3), 185–210 (1983)
86. I. Press, R.F. Ganey, M.P. Malone, Satisfied patients can spell financial Well-being. *Healthcare financial management: journal of the Healthcare Financial Management Association**45***(2), 34–36 (1991)., 38, 40-32
87. S. Purbey, K. Mukherjee, C. Bhar, Performance measurement system for healthcare processes. Int. J. Product. Perform. Manag. **56**(3), 241–251 (2007)
88. Reason, P., & Bradbury, H. (2001). *Handbook of action research: Participative inquiry and practice*: Sage
89. C. Schoen, R. Osborn, D. Squires, M. Doty, P. Rasmussen, R. Pierson, S. Applebaum, A survey of primary care doctors in ten countries shows progress in use of health information technology, less in other areas. Health Aff. **31**(12), 2805–2816 (2012)
90. C.L. Schur, L.A. Albers, M.L. Berk, Health care use by Hispanic adults: Financial vs. non-financial determinants. Health Care Financ. Rev. **17**(2), 71–88 (1994)
91. S.M. Shortell, J.P. LoGerfo, Hospital medical staff organization and quality of care: Results for myocardial infarction and appendectomy. Med. Care **19**, 1041–1055 (1981)
92. D.M. Steinwachs, R.G. Hughes, *Health Services Research: Scope and Significance* (An Evidence-Based Handbook for Nurses, Patient Safety and Quality, 2008), pp. 08–0043
93. E.R. Stinson, D.A. Mueller, Survey of health professionals' information habits and needs: Conducted through personal interviews. JAMA **243**(2), 140–143 (1980)
94. A. Stoop, M. Berg, Integrating quantitative and qualitative methods in patient care information system evaluation: Guidance for the organizational decision maker. Methods Inf. Med. **42**(4), 458–462 (2003)
95. E. G. Sukoharsono, C. SE. Where Are You, Papi? Asking Mami indoubt with Papi: An Imaginary Dialogue on Critical Ethno-Accounting Research
96. A. Sunyaev, A. Kaletsch, H. Krcmar. *Comparative evaluation of google health API vs. Microsoft healthvault API.* Paper presented at the proceedings of the third international conference on health informatics (HealthInf 2010) (2010)
97. J.R. Swisher, S.H. Jacobson, J.B. Jun, O. Balci, Modeling and analyzing a physician clinic environment using discrete-event (visual) simulation. Comput. Oper. Res. **28**(2), 105–125 (2001)
98. H.C. Szeto, R.K. Coleman, P. Gholami, B.B. Hoffman, M.K. Goldstein, Accuracy of computerized outpatient diagnoses in a veterans affairs general medicine clinic. Am. J. Manag. Care **8**(1), 37–43 (2002)

99. G. Thomas, A typology for the case study in social science following a review of definition, discourse, and structure. Qual. Inq. **17**(6), 511–521 (2011)
100. S. Thomas, R. GLYNNE-JONES, I. CHAIT, Is it worth the wait? A survey of patients' satisfaction with an oncology outpatient clinic. Eur. J. Cancer Care **6**(1), 50–58 (1997)
101. S. Tsumoto, Knowledge discovery in clinical databases and evaluation of discovered knowledge in outpatient clinic. Inf. Sci. **124**(1), 125–137 (2000)
102. K. Van Peursem, M. Prat, S. Lawrence, Health management performance: A review of measures and indicators. Account. Audit. Account. J. **8**(5), 34–70 (1995)
103. J.A. Vennix, J.W. Gubbels, Knowledge elicitation in conceptual model building: A case study in modeling a regional Dutch health care system. Eur. J. Oper. Res. **59**(1), 85–101 (1992)
104. M.M. Wagner, W.R. Hogan, The accuracy of medication data in an outpatient electronic medical record. J. Am. Med. Inform. Assoc. **3**(3), 234–244 (1996)
105. B.B. Wang, T.T. Wan, D.E. Burke, G.J. Bazzoli, B.Y. Lin, Factors influencing health information system adoption in American hospitals. Health Care Manag. Rev. **30**(1), 44–51 (2005)
106. B.B. Wang, T.T. Wan, J. Clement, J. Begun, Managed care, vertical integration strategies and hospital performance. Health Care Manag. Sci. **4**(3), 181–191 (2001)
107. S.-W. Wang, W.-H. Chen, C.-S. Ong, L. Liu, & Y.-W. Chuang. *RFID application in hospitals: a case study on a demonstration RFID project in a Taiwan hospital.* Paper presented at the system sciences, 2006. HICSS'06. Proceedings of the 39th annual Hawaii international conference on (2006)
108. R. Wilton, J. M. McCoy. An outpatient clinic information system based on distributed database technology. Paper presented at the Proceedings/the... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care (1989)
109. J.D. Wisner, S.E. Fawcett, Linking firm strategy to operating decisions through performance measurement. Prod. Invent. Manag. J. **32**(3), 5–11 (1991)
110. H.L. Wu, Y.Y. Liu, C.J. Dong, K. Wang, Multiple factors integration based text alignment for medical database. Adv. Mater. Res. **756**, 1648–1651 (2013)
111. E. Xu, M. Wermus, D.B. Bauman, Development of an integrated medical supply information system. Enterprise Inform. Syst. **5**(3), 385–399 (2011)
112. F. Yeşilada, E. Direktör, Health care service quality: A comparison of public and private hospitals. Afr. J. Bus. Manag. **4**(6), 962–971 (2010)
113. F. Youssef, D. Nel, T. Bovaird, Service quality in NHS hospitals. J. Manag. Med. **9**(1), 66–74 (1995)
114. B. Zeng, A. Turkcan, J. Lin, M. Lawley, Clinic scheduling models with overbooking for patients with heterogeneous no-show probabilities. Ann. Oper. Res. **178**(1), 121–144 (2010)

Chapter 3
Smart City: An Intelligent Automated Mode of Transport Using Shortest Time of Travel Using Big Data

Mashrin Srivastava, Suvarna Saumya, Maheswari Raja, and Mohana Natarajan

3.1 Introduction

With the focus shifting to revolutionizing the way of life by introducing 'smart' technology to analyse and predict needs or reduce the man power required to complete jobs and the onset of emerging technology in the field of big data analytics and cloud computing in this age of technology, the area of focus in our proposal is automated transportation. The proposal is to use a modified Dijkstra's algorithm to calculate the shortest time to reach the destination instead of the shortest path. The simple explanation for choosing this is that in today's fast paced world, the constraint of time is more pressing as compared to distance. People want results faster. Time is money. With everyone scrambling to squeeze in a few extra seconds, they want to save time while travelling from one place to another. The fastest path is however determined relative to the others on the road. If there is a shortcut to a certain place, then one can hope to reach that place faster by using that shortcut. However if a large number of drivers choose the same path, that path becomes a bottleneck to reaching the destination. Another constraint is the condition of the roads and the speed limits on the roads. If there is a small dirt path and a concrete path, it is an obvious conclusion that the speed will be faster on the concrete roads other than the dirt path. Where automated cars are concerned, it is needed to take the surroundings like speed bumps, construction work, etc. into consideration [1, 2]. These factors play a pivotal role while choosing a suitable path that will help reach the destination in the shortest time possible [3].

M. Srivastava · S. Saumya · M. Raja (✉) · M. Natarajan
Vellore Institute of Technology, Chennai, India
e-mail: mashrin.msrivastava2014@vit.ac.in; suvarna.saumyacjyoti2014@vit.ac.in; maheswari.r@vit.ac.in; mohana.n@vit.ac.in

S. Paul et al. (eds.), *Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-77558-2_3

3.1.1 Background and Motivation

The latest advancement in technological innovation raising to a deployment of smart city targeting imminent in smart vehicle mobility with Artificial Intelligence Centric, Machine Learning Centric, IoT enabled automation supporting Industry 4.0 as well. All ways of remote automation is achieved through IoT which ensures exceptional connectivity planes amongst the users and vehicles. The existence of flexibility in restructuring the merchandise logistics and transport optimization using various digital technologies such as remote/real automation, connected car techniques, big data analytics technologies and Industry 4.0 forms the strength of tomorrow's transportation and logistics market productivity assuring smart transport system [4]. The diverse cohorts of Connected and Automated vehicles (CAV) with numerous echelons of system automation along with human-driven vehicles (HDV) make us to extend and achieve extra resourceful and safer transportation systems [5]. Hosting of autonomous buses empowers optimization with reduction of fleets and cost acquired by the vehicle. The prototypical construction method anticipated castoff to deliver bus transit with effective guidance counting autonomous buses on its integrational bus fleets configuration and distribution. Henceforward, this work provides the outline of the research practice employed for an automated based Society of Automotive Engineers level five mode of transport expending a modified Dijkstra's algorithm integrating ant colony optimization, big data analytics and cloud computing. This integration helps the driver to reach the destination based on the traffic pertaining to the shortest time possible instead of the usual shortest distance algorithms.

3.2 Literature Survey

3.2.1 Automated Transport

Experiments in the field of automated cars started in the 1920s gained momentum in 1950s and is still improving with projects like Navlab from Carnegie Mellon University, Vislab from University of Parma, Eureka Prometheus Project from Bundeswehr University Munich, etc. Automobile companies like Tesla motors, Mercedes-Benz, Renault, Toyota, Bosch, Nissan, etc., are amongst the leaders in prominent projects undertaken [6, 7]. The most famous project is Google Self-Driving car undertaken by Google X.

3.2.2 Classification of Autonomous Cars

The standardization body for automotive-SAE (Society of Automotive Engineers) has a classification system consisting of six different levels based on the involvement of the driver based on his attentiveness and intervention that is required to reach the destination safely more than the capabilities of the vehicles [8].

Level 0: The automated system may alert the driver but it does not control the vehicle.

Level 1: Features such as ACC (Adaptive cruise control), Parking assistance with automated steering and LKA (Lane Keeping Assistance) in some combination requires the driver to be in a state ready to take control anytime.

Level 2: The driver must take control of detecting objects and events if the automated system fails to respond properly. When the driver takes over, all activities like accelerating, braking and steering must be suspended by the autonomous system and control must be promptly passed to the driver.

Level 3: The driver can afford to be inattentive towards the driving tasks only in known limited environments like freeways.

Level 4: The driver needs to control the vehicle in severe conditions like bad weather apart from which the automated system can take control of the vehicle without requiring the driver's attention.

Level 5: The driver is only required to start the system and set the destination, and the automated system can take control of driving in any location provided it is legal to drive.

3.2.3 Ant Colony Optimization

Ant colony Optimization algorithms is a part of swarm intelligent methods that use probabilistic techniques to solve problems. The basic principle comes from the food gather techniques of an ant colony [9]. The initial movement of ants is random; however, they leave a trail of pheromones on their way to the colony. If any of the other ants moving at random happen to see this food trail, they follow it and in turn leave pheromones increasing the strength of the pheromone trail which makes it more likely for the ants moving randomly to be attracted to it. With time however the pheromone trails evaporate and hence the longer the path from the food to the colony is, the more the chances of the ants not following the path due to disappearance of the pheromone trail, but in comparison, the shorter paths have more concentrated pheromones thus having a much higher probability of ants following that path and reinforcing the pheromones till finally all ants are following one path. The different variations of the ant colony optimization include the 'Min–max ant system', 'Elitist ant system', 'Rank-based ant system', etc. The applications include vehicle routing, scheduling, assignments, set, image processing, protein folding, data mining, etc.

3.2.4 Dijkstra's Algorithm

For a given weighted graph input, Dijkstra's algorithm can be used to find the shortest path between vertices in the graph when the condition that all weights are non-negative is satisfied. Originally the Dijkstra's algorithm found the shortest path between two vertices, but a more popular variation that is widely used is for finding the shortest path from one vertex to all other vertices in the graph [10]. Dijkstra's algorithm has many real-world applications that include finding the shortest path from one city to another using the driving distance as the weights in the graph and the nodes as the different cities. One major pitfall is the exclusion of the concept of quantity and time i.e. if there are many cars and the corresponding effects of the time taken to reach the destination [11, 12]. Using the concepts of Ant Colony Optimization (ACO) and Dijkstra's algorithm, a modified version can be tailored to suit this need [13].

3.2.5 GPS (Global Positioning System)

Irrespective of the weather conditions, be it rain, hail or snow if there are four satellites that have access to direct view of the GPS receiver, the current time and location are provided to the person by the navigation system based in space. The GPS system was created by the USA in addition to which various systems were in use and development including Global Navigation Satellite System (GLONASS) from Russia, BeiDou Navigation Satellite System from China, Quasi-Zenith Satellite System from Japan and Indian Regional Navigation Satellite System from India [14].

3.2.6 LIDAR (Light Detection and Ranging)

A laser is popularly used in fields like geology, geomatics, geomorphology, etc., to illuminate the target in order to measure distance [15].

3.2.7 Odometry

It involves the measurement of the current position relative to the position in which it started with the help of motion sensors [16]. The sensitivity to errors is high because the measurement of position involves the integration of velocity over time. Object recognition and event recognition is most needed in the implementation of the proposal.

3.2.8 Computer Vision

For a given weighted graph input, Dijkstra's algorithm can be used to find the shortest path between vertices in the graph when the condition that all weights are non-negative is satisfied. Originally the Dijkstra's algorithm found the shortest path between two vertices but a more popular variation that is widely used is for finding the shortest path from one vertex to all other vertices in the graph [10]. Dijkstra's algorithm has many real-world applications that include finding the shortest path from one city to another using the driving distance as the weights in the graph and the nodes as the different cities. One major pitfall is the exclusion of the concept of quantity and time i.e. if there are many cars and the corresponding effects of the time taken to reach the destination [11, 12]. Using the concepts of Ant Colony Optimization (ACO) and Dijkstra's algorithm, a modified version can be tailored to suit this need [13].

3.2.9 Ad Hoc Network

Ad hoc network is a type of wireless network that is decentralized, and it is free from the constraint of existing infrastructure because it relies on each node in the network to behave like a router and forwards data to the other nodes thus dynamically deciding the next node based on the available connectivity [17]. Mobile ad hoc networks (MANET) comprise mobile devices that follow the principles of ad hoc networks [18]. Vehicular ad hoc networks (VANETS) comprise the application of MANET to the domain of vehicles where the mobile devices are the vehicles themselves making them the main component of intelligent transport systems (ITS) with uses in electronic brake lights, traffic information systems and platooning.

3.2.10 Big Data

Big data refers to copious amounts of real-time data that has the characteristics of a large amount of data that can even extend to petabytes of data (volume), different types of data including audio, text, images, videos, etc., (variety), varying speed at which the data is collected (velocity) and the difference in the quality of data captured (veracity) making it too large and complex for traditional data processing applications to suffice. Application of big data to decision making gives better accuracy in results and helps make more informed decisions for more efficiency and reduced risks and cost [19]. Big data analytics involves combing through the vast amount of data to find unknown correlations and hidden patterns to provide information that could help in the decision making process. In this project the intention is to harness this power of big data analytics in an endeavour to make the vehicle secure by using it to find the shortest path, to detect any obstacles through vision computing and image analysis and to provide better response dynamically to the vast amounts of real-time data generated so that decisions can be made spontaneously [20]. Predictive models can be used to determine the times and places for maximum traffic enabling the system to cope with the problems brought with it.

3.2.11 Cloud Computing

Cloud computing involves the sharing of computer resources and other services like applications, virtual machines, etc., on demand to a variety of devices. The advantage of cloud computing includes cost saving in terms of hardware equipment, agility, increased security due to decentralized data, independence of location and device, reliability, scalability, etc. The main advantage includes the provision of resources rapidly and with minimal amount of managerial effort [21]. The ability to provide prompt services in terms of expansions and the huge amount of computational power that is available due to sharing of resources makes cloud computing an essential part of the project.

3.3 Motivation

The motivation behind this started with the limitations of the autonomous cars as well as the limitations for Dijkstra's algorithm. For the autonomous cars, testing in heavy snow and rain has not been carried out due to safety concerns. Even with the use of LIDAR, it becomes difficult to distinguish light debris and garbage causing the car to veer unnaturally without any real threat and inability to distinguish police officers signalling cars to stop. The first known incident that proved fatal involved Tesla model S, an electric car from Tesla motors when an 18 wheel trailer tractor made a left turn just in front of the Tesla car and the car failed to stop on the non-controlled access highway [22–24]. These specified limitations can be overcome by incorporating appropriate weather forecasting and weather mapping techniques. The adverse weather conditions can be detected and necessary precautions like slowing down the speed in heavy rains to prevent skidding can be taken forward. The accidents involving Google's and Tesla's car were due to human error. This occasion does not arise when all cars are being controlled [25, 26]. For the Dijkstra's algorithm, the major disadvantage is that the algorithm does a blind search thereby consuming a lot of time and the wastage of necessary resources. Another disadvantage is that it cannot handle negative edges, which leads to acyclic graphs and most often cannot obtain the right shortest path. But the biggest problem with the algorithm is that though it is widely used, we believe that it is not an algorithm that should be used for path navigation. The traditional navigation applications made use of this algorithm initially given that it is well known and works well for finding the shortest path. Since in the past, traffic was not a major issue because of lesser number of vehicles on the road, this algorithm performed fairly well. But this is not the case now. There are a lot of vehicles on the roads now, leading to heavy traffic. We have all observed in day to day life, that the shortest path now is no longer the path which will take the minimum time to reach somewhere. This was the motivation for proposing a modification to the Dijkstra's algorithm so that the heuristic takes care of this fact and suggests the path which takes the least time rather than the least distance as the value for time is becoming more and more important in the lives of people given the fast moving world and our lifestyles.

3.4 Problem Statement

Given a weighted graph of the road network, the objective is to find the path that takes the least amount of time to reach from any given location A to location B, where each location is represented as the vertex in the graph.

3.5 Proposed Work Description

The complete path flow of the proposed work is shown in Fig. 3.1. The major components of flow diagram are start, choose destination, shortest travel time path set, automated transport, destination arrive and stop.

- *Start*: The user is logged into the system as soon as the car is turned on. Choose destination: The source and destination is chosen based on the current position as identified by the GPS system and the input by the user, respectively.
- *Shortest travel time path set*: Using the proposed modified algorithm, based on all the users currently logged onto the system and their current position and destination, the route that takes the shortest time to travel is set.
- *Automated transport*: Using the grid based system where the inaccessible roads are blocked out by giving a very high value and the weight is taken in accordance to the speed the car can achieve on the particular road, the user is transported from source to destination.
- *Arrive at destination*: With minimal sensors, this automated mode transport uses big data analytics and cloud computing to ferry people from one place to another safely.
- *Stop*: Once the car is stopped, the user is logged off the system.

The first part of the prototype consists of the modified Dijkstra's algorithm in which the shortest path is determined based on the time taken and not on the distance between the source and destination. Analysing is required to find out the number of people who had own car and statistically how many people would want to go to the same place at the same time in a particular place. If the shortest route is to take the I-

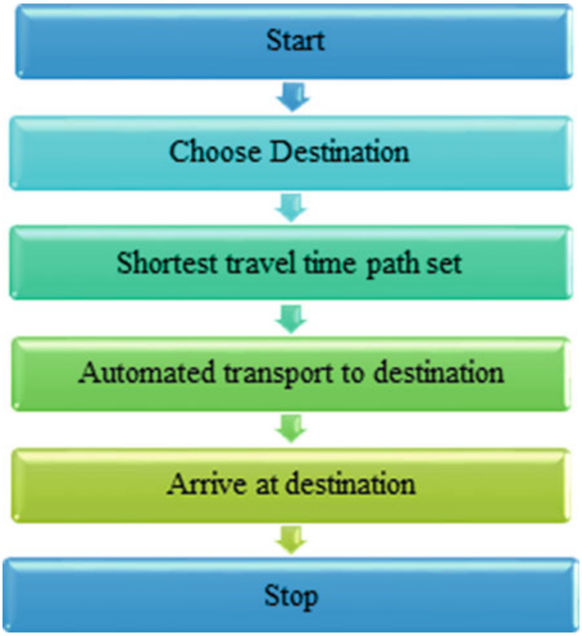


Fig. 3.1 Flowchart of the proposed work

16 and turn north towards the bridge, and all 100 people take the I-16 and turn north, then imagine the traffic jam a vehicle would be stuck in with the agonizing honking and bumper to bumper traffic. The whole ground can be represented as a grid or a matrix. A path can then be plotted based on the parts of the grid that are free. For example, if there is a pothole then that part of the grid can be made as inaccessible and will not show as a viable route that can be taken. Similar pothole checking and path plotting is made to all the places that are on either side of the road. In Fig. 3.2a and b the grey area represents a concrete road on the highway, therefore the weight assigned is minimal. The yellow area represents a tar road near residential areas, therefore the weight assigned is 10. The black areas are the areas that should not be accessed. We can either give a very large weight representing infinity.

With the offset of the GPS system and the popular google maps, the dependence on maps or memorizing routes has decreased leading to the ease in travel due to assurance of arrival at the mentioned destination. The project uses this concept to ensure accuracy in reaching the destination without needing to rely on maps or roads. A number of safety features will be provided including the application of vision computing to pre-detect objects and potentially dangerous situations.

Steps in the Algorithm for the Modified Dijkstra

- Start.
- Initialize the map.
- The values according to the speed in the particular area are set. The inaccessible zones and areas of roadwork are initialized with infinity value so that the areas are never chosen.

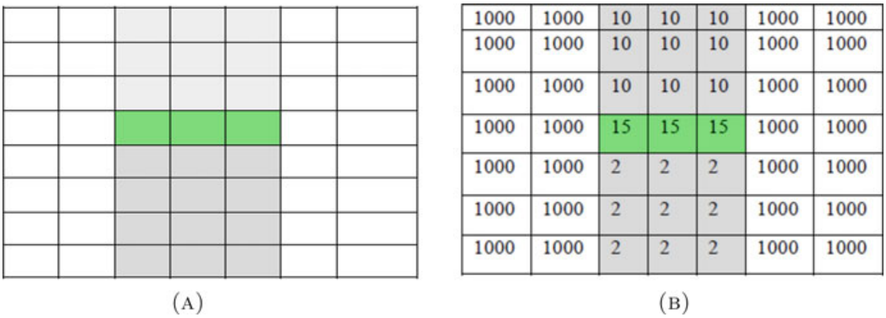


Fig. 3.2 Sample maps. (**a**) Sample map (blank parts). (**b**) Sample map with weights

- For the first driver is assigned the shortest path available to it and the route congestion for that path is incremented.
- For all the subsequent drivers, the most optimal path according to the modified shortest path algorithm is calculated considering the present congestion of the paths
- Optimal path is calculated by increasing the initial weight of the path by some factor of congestion, Weight = Weight + k*congestion.
- The destination is reached by providing the most optimal path using the modified algorithm.
- The data is updated dynamically to include the changes in traffic due to unforeseeable factors.
- Stop.

The time saved by each driver when the modified algorithm is used to calculate the shortest path to the destination instead of Dijkstra's algorithm is calculated for map of sizes 4×4, 6×6, 8×8 and 10×10. Figure 3.3 shows the values plotted and the corresponding curve fitting for the plotted value. The slope of the curve increases with increase in the size of the matrix showing the increase in the time saved as the size of the map increases. In a real-world scenario, the size of the matrix would be enormous as the roads and surrounding would be divided into multiple grids decreasing the size of the grids for more accuracy.

3.6 Implementation of Prototype

The model of the proposal was made using Arduino and the functioning was tested along with the application of the modified Dijkstra's algorithm. The path to be followed is traced using a vector, and the model robot follows the path that is given. In the proposal, the same will happen with the help of Google maps API. Using GPS the starting point of the user can be located and when the user enters the destination, the vectors to reach the destination is fetched by the Google maps API in accordance

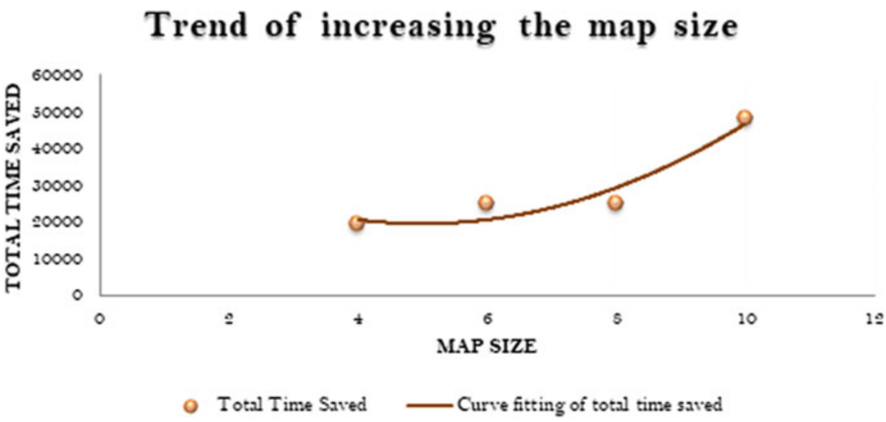


Fig. 3.3 Graph showing the total time saved as the map size increases

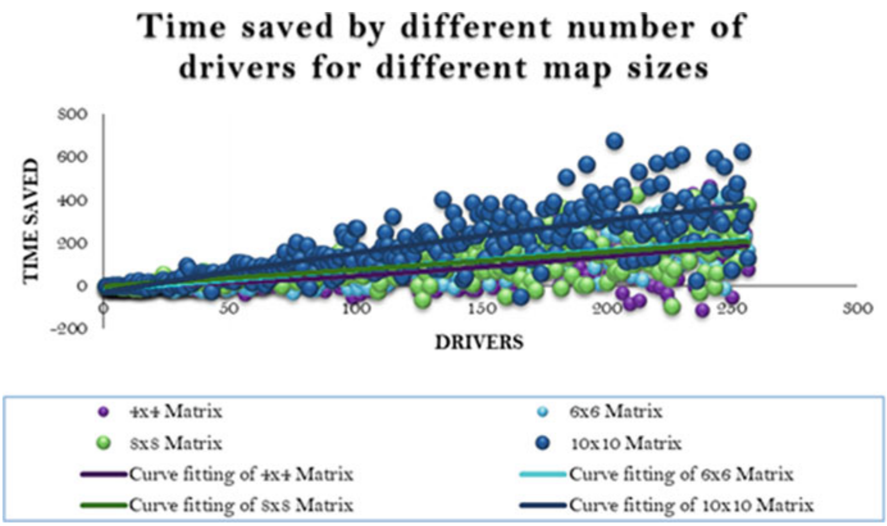
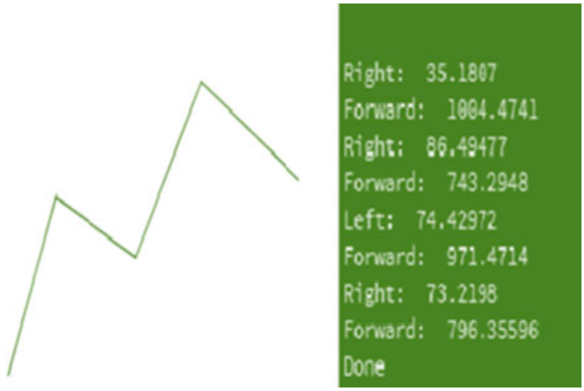


Fig. 3.4 Graph plot for the time saved by each driver

to the modified algorithm. The model moves along the plotted path as shown in Fig. 3.4. The scale is specified as the actual movement with respect to the length of the plotted vector. The model then moves along the given path by calculating the length by multiplying the plotted length with the factor of the scale. The same principle is applied using the Google API. Figure 3.4 also shows the actual length moved by the model. Figure 3.5 shows the movement of the model in progress. In Figure 3.6 are the screenshots that show the modified algorithm for the matrix of size 5×5 and the paths allotted for the different drivers having different source and destination.

Fig. 3.5 Time saved in a particular case

Fig. 3.6 Vector plot and movement of the model



3.6.1 Mathematical Model

Let d_i, t_i and V_i be the distance, time and average velocity for the path 'i' given in Eqs. (3.1) and (3.2).

$$\text{Average Velocity } V_i = d_i/t_i \quad (3.1)$$

$$\text{Optimal Route } = Max\{V_1, V_2, , V_i\} \quad (3.2)$$

The trial runs performed for a number of matrices of varying sizes and the time saved by each of the driver are plotted. Figures 3.4 and 3.5 represent the time saved by each driver and time saved in a particular case. As the map size increases, we

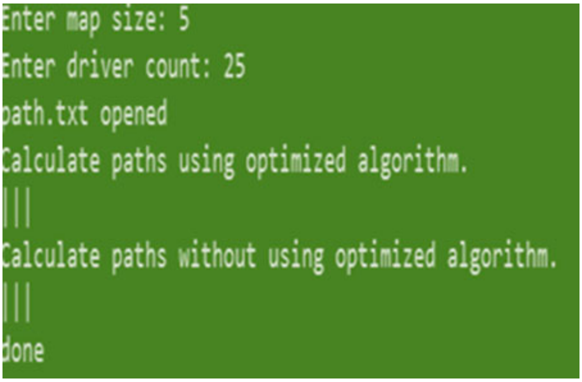


Fig. 3.7 Screenshot of application of the algorithm

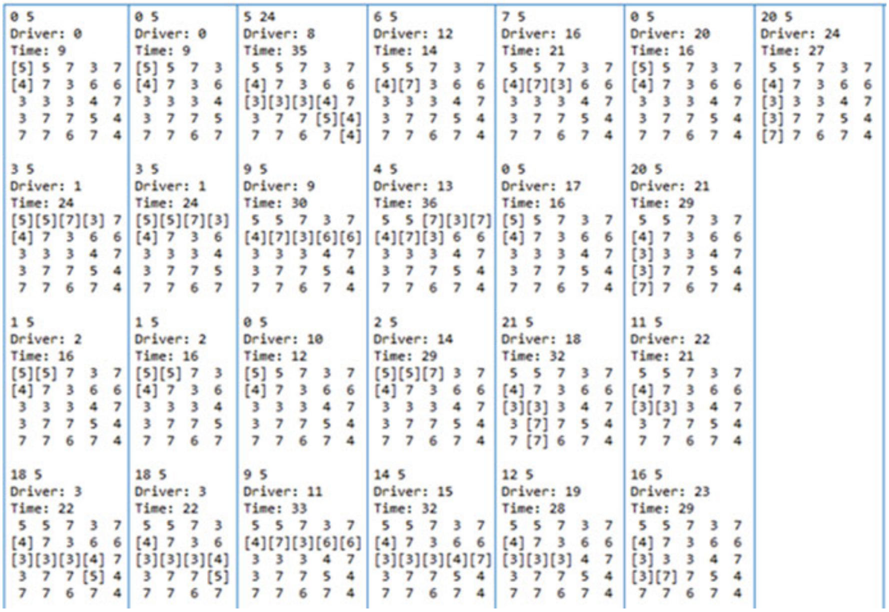


Fig. 3.8 The path selection using the modified algorithm

can observe that the time saved for each driver is increased. This Fig. 3.6 shows that with the increase in map size that the real-world scenario will have, the time saved will increase thus proving the efficiency of the modified algorithm and ensuring high performing automated cars.

3.6.2 Collision Avoidance

The most important issue to be addressed is that of collision avoidance. There are two types of collision—static and dynamic. Although in the proposal there is no possibility of collision but taking into account the errors and malfunctioning that could happen, the possibility of errors and the corresponding methods for correction have to be taken into consideration. For static collisions in case of roadwork or inaccessible paths that can be predefined before the vehicle reaches the area, since the map is updated dynamically, they can be assigned a very high value that forbids the movement to those areas. For the dynamic collisions or the collisions that may occur between the vehicles on the road, we can use vision computing methods for analysis of the movement of the vehicles in the vicinity. The results then can be communicated from vehicle to vehicle using the principles of VANET. In case of manual rerouting of traffic, a specialized VANET signal can be used to communicate the message to the vehicles thus successfully rerouting traffic in case of emergencies and even in the case of ambulance travels. The future work would be to build a prototype of the model and go over the safety features in the case of a failure with a fine comb.

A prototype of the proposed work is designed and implemented using Arduino and appropriate hardware components.

Performance analysis has been made by constructing the code to calculate the paths using optimized algorithm and paths without using optimized algorithm with map size of 5 and driver count of 25. The obtained result is depicted in Fig. 3.7 as screenshot of application of the algorithm.

The traces of varying values of the car driving with respect to number of drivers and time stamp is simulated and tracked appropriately. Figure 3.8 shows the output of the path selection using the modified Dijkstra algorithm.

3.7 Conclusion and Future Work

Thus, a prototype is proposed and analysed for the mode of transport using a modified Dijkstra's algorithm. Deriving the concepts of Ant Colony Optimization, big data analytics and cloud computing, a novel technique to arrive at the destination based on the traffic and the number of drivers on the road, in the shortest time possible instead of the usual shortest distance algorithms is simulated. Appropriate analysis with respect to time frame is performed and presented here to validate the state of the art. The graph and screenshots remain the proof of concept supporting the correctness of modified Dijkstra algorithm for shortest time. The future work would be to rebuild a prototype of the model and go over the safety features in the case of a failure with a fine comb.

References

1. F. Nashashib, L. Bouraoui, A cooperative personal automated transport system *"A CityMobil Demonstration in Rocquencourt"*, in *2012 12th International Conference on Control, Automation, Robotics and Vision Guang zhou, China, (ICARCV 2012)* (2012), pp. 644–649
2. L. Bouraoui, F. Charlot, C. Holguin, F. Nashashibi, M. Parent, P. Resende, An on-demand personal automated transport system: the CityMobil demonstration in La Rochelle, in *Intelligent Vehicles Symposium (IV), June 2011* (2011), pp. 1086–1091, pp. 5–9
3. M. Parent, Cybercars: A solution for urban transport, in *CODATU Conference* (2004)
4. S. Hamdar, A. Talebpour, R. Bertini, Traffic and granular flow: the role of data and technology in the understanding of particle dynamics. J. Intell. Transp. Syst. Technol. Plann. Oper. **24**(6), 535–538 (2020)
5. A. Nikitas, K. Michalakopoulou, E. Tchouamou Njoya, D. Karampatzakis, Artificial intelligence, transport and the smart city: definitions and dimensions of a new mobility era. Sustainability **12**, 1–19 (2020)
6. J. Choi, Y. Khaled, M. Tsukada, T. Ernst, IPv6 support for VANET with geographical routing, in *8th International Conference on Intelligent Transport System Telecommunications (ITST 2008), Phuket* (2008)
7. J. Xie, J. Xie, F. Nashashibi, M. Parent, O. Favrot, A real-time robust global localization for autonomous mobile robots in large environments, in *ICARCV 2010* (2010), pp. 1397–1402
8. T. Meyerowitz, C. Pinello, A. Sangiovanni-Vincentelli, A tool for describing and evaluating hierarchical real-time bus scheduling policies, ACM, Anaheim, California, USA, in *Proceedings of the 40th Annual Design Automation Conference* (2003), pp. 312–317
9. L. Zuo, L. Shu, C. Zhu, T. Hara, A multi-objective optimization scheduling method based on the ant colony algorithm in cloud computing, in *IEEE Special Section on Big Data Services And Computational Intelligence for Industrial Systems*, vol. 3 (2015), pp. 2687–2699
10. X. Cheng, L. Fang, L. Yang, S. Cui, Mobile big data: the fuel for data-driven wireless. IEEE Internet Things J. **4**(5), 1489–1516 (2017)
11. M.A. Alsheikh, D. Niyato, S. Lin, H.-P. Tan, Z. Han, Mobile big data analytics using deep learning and apache spark. IEEE Netw. **30**(3), 22–29 (2016)
12. J.L. Toole, C. Herrera-Yaqe, C.M. Schneider, M.C. Gonzlez, Coupling human mobility and social ties. J. R. Soc. Interface **12**(105), 1–9 (2015)
13. H. Dong, Traffic zone division based on big data from mobile phone base stations. Transp. Res. C Emerg. Technol. **58**, 278–291 (2015)
14. S. Sheeba Rani Gnanamalar, R. Maheswari, B. Sharmila, V. Gomathy, IoT driven vehicle license plate extraction approach. Int. J. Eng. Technol. (UAE) **7**(2), 457–459 (2018)
15. S. Massobrio, A. Nesmachnow, A. Tchernykh, A. Avetisyan, G. Radchenko, Towards a cloud computing paradigm for big data analysis in smart cities. Program. Comput. Softw. **44**(3), 181–189 (2018)
16. S. Oh, Y.J. Byon, H. Yeo, Improvement of search strategy with K-nearest neighbours approach for traffic state prediction. IEEE Trans. Intell. Transp. Syst. **17**(4), 1146–1156 (2016)
17. M. Peng, K. Zhang, J. Jiang, J. Wang, W. Wang, Energy-efficient resource assignment and power allocation in heterogeneous cloud radio access networks. IEEE Trans. Veh. Technol. **64**(11), 5275–5287 (2015)
18. B. Bangerter, S. Talwar, R. Arefi, K. Stewart, Networks and devices for the 5G era. IEEE Commun. Mag. **52**(2), 90–96 (2014)
19. Q. Shi, M. Abdel-Aty, Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. Transp. Res. C Emerg. Technol. **58**, 380–394 (2015)
20. H. Song, H. Liang, H. Li et al., Vision-based vehicle detection and counting system using deep learning in highway scenes. Eur. Transp. Res. Rev. **11**, 51–55 (2019)
21. C. Zhu, X. Li, V.C.M. Leung, X. Hu, L.T. Yang, Job scheduling for cloud computing integrated with wireless sensor network, in *Proc. IEEE 6th Int. Conf. Cloud Computing. Technol. Sci. (CloudCom)* (2014), pp. 62–69

22. F. Farahnakian et al., Using ant colony system to consolidate VMs for green cloud computing. IEEE Trans. Services Comput. **8**(2), 187–198 (2015)
23. B. Zhang, X. Wan, J. Luo, X. Shen, A nearly optimal packet scheduling algorithm for input queued switches with deadline guarantees. IEEE Trans. Comput. **64**(6), 1548–1563 (2015)
24. Z. Tang, L. Jiang, J. Zhou, K. Li, K. Li, A self-adaptive scheduling algorithm for reduce start time. Future Generat. Comput. Syst. **43–44**(3), 51–60 (2015)
25. Y. Chen, A. Zhang, Z. Tan, Complexity and approximation of single machine scheduling with an operator non-availability period to minimize total completion time. Inf. Sci. **25**(1), 150–163 (2015)
26. Y. Zha, J. Yang, Task scheduling in cloud computing based on improved ant colony optimization. Comput. Eng. Des. **34**(5), 1716–1719 (2013)

Chapter 4
Context Awareness for Healthcare Service Delivery with Intelligent Sensors

Shikha Singhal, Adwitiya Sinha, and Buddha Singh

4.1 Overview

Our research surveys through several application-based challenges and recent trends. It includes a complete systematical introduction of fundamental concepts of e-healthcare systems and the importance of contextual sensors with user mobility. Later, more complex and advanced topics including a case study are included to develop a research initiative required for implementing secure framework for healthcare service delivery.

4.1.1 Healthcare Service Delivery

Traditionally, providing an unwavering healthcare service for elders, loved ones, and especially the person with disabilities had become a critical issue that exacerbates the situation faced by family members. Nowadays, people usually having busy schedules do not have much time, as a result of which health issues are ignored and therefore, eventually tend to avoid consultation with doctors regularly. However, certain critical health issues, like cardiovascular and chronic diseases, cannot be ignored and hence require attention consistently.

S. Singhal · A. Sinha (✉)
Department of Computer Science & Engineering and Information Technology, Jaypee Institute of Information Technology, Noida, Uttar Pradesh, India

B. Singh
School of Computer & Systems Sciences, Jawaharlal Nehru University, New Delhi, India

S. Paul et al. (eds.), *Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-77558-2_4

Rapidly, the world population is increasing day by day along with healthcare costs that are raising forcibly. On the comparison with traditional healthcare, we normally do not often lean towards digitized equipment to check our health issues regularly and timely. Therefore, efforts are being made towards ubiquitous technology in performing medical observation in a pervasive manner. This would help in transforming the healthcare system that allows continuous monitoring of inhabitants without hospitalization. The rapid development in the technology area provides us a new way of monitoring health through using mobile Internet, cloud computing technology, sensor technology, and intelligence system. Now the goal is to develop a pervasive application, or system for delivering a mobile healthcare service to customers anywhere, anytime with the help of the above technologies [1–3].

Deployment of intelligent sensors pervasively has revolutionized the delivery of sustainable healthcare services over time and space. Owing to technological advancements in electronics, sensors are embedded within the scenario we dwell in, thereby resulting in better sampling and availability of information anytime and anywhere. This next-generation infrastructure has also developed high-end mechanisms for accurate medical information mining, storage, and retrieval. This helps in building, evaluating, and presenting an improved platform for transforming the traditional healthcare systems towards innovative heights using intelligent sensor technology.

Environmental monitoring is one of the most significant and widespread applications of intelligent sensors that allows a cluster of sensing devices to monitor various environmental parameters and conditions over a stipulated period. Such healthcare services for monitoring health-related issues of patients together with significantly reducing the healthcare expenditures are incurred in present healthcare systems. These sensors can also be used to provide a more useful, collaborative, and intelligent living environment for human beings. For instance, ambient intelligence is growing in the form of a future vision of intelligent sensing and managing the delivery of medical information worldwide. This technology offers a paradigm that is capable of monitoring health status periodically as well as continuously. It can even assist in diagnosing health conditions or even can get indulged in communicating with the patient regarding lifestyle or food habits for maintaining a sound health. Ambient technology allows to remotely contact with medicinal experts for regular check-ups or in case of a medical emergency.

The pervasive availability of sensory data enhanced with context information addresses existing healthcare problems with a promising future. Context-aware services imparted by intelligent sensing electronics results in advanced healthcare delivery to patients by eradicating location, temporal, financial, and other resources limitations without compromising service coverage, timely delivery, and quality of assistance. Such advancements broadly include prevention from viral infections, regular health check-ups, healthcare maintenance with customized diets, etc. Several real-time applications extend support in making advanced healthcare services feasible. This largely involves smart homes and hospitals, advanced cloud services,

Fig. 4.1 Context-aware modeling for healthcare service delivery with intelligent sensors

mobile devices and applications, autonomous systems, etc. The intelligent healthcare system mainly includes the following significant characteristics in principal:

- Short-term healthcare monitoring
- Intelligent homing
- Long-term nursing
- Tracking health improvements
- Personalized medical supervision.
- Incidence detection
- Risk management in disasters
- Emergency medical intervention
- Transportation facilities
- Treatment updates

These features promise an effective solution with reliable access to widely distributed healthcare services, including biomedical information sharing with remote medical and surgical units with pervasive availability to attend for services or enquires.

Figure 4.1 explains the whole scenario, how mHealth services are provided by healthcare professionals at anytime and anywhere. Here, all the recorded information of patient is stored on a server that can only be retrieved by doctors and healthcare professionals for providing better treatment on time and anywhere.

4.1.2 Significance of Context Information

The word "context" refers to surrounding information for a given situation, which can be time, date, or any object of interest. It helps users to develop a consciousness of behavior with an awareness of the environment. Representation of context to communicate with the system in a more efficient manner is known as contextual information. In other words, contextual information is any relevant information regarding the system and its users so that system can deliver personalized service as per user requirements. In the healthcare service delivery system, we need context information to deal with different users. How can be context information helps us! This question arises in our mind. In the future, the intelligent hospital will be built that would introduce new technologies, new architectures so that the implementation of the new system will be secure and reliable. The main concern is to identify, evaluate, and implement a new system or services without any hazards which provide good communication between healthcare professionals and co-operate with users, doctors, and healthcare service providers (HSPs) [4].

Dedicated Use of Context According to the researchers, we can use context in three main cases:

- To represent relevant information and service of the systems for the user
- To execute new service accurately and properly
- Tag relevant context information for lateral usage

Objects for the Representation of Context For analyzing the healthcare contexts, it is possibly necessary to explain the distinct contexts in detail. To represent the context in a synthesized way, we split the context items into three main classes: (i) environment, (ii) people, and (iii) activities.

Context-Aware Computing It refers to the building of applications that is capable of adapting certain changes in the surrounding (people, environment, activities) and responds according to the usage of the context item in that situation [5].

Context-awareness computing is still a challenging concept in healthcare services. The three major challenges are [6]:

- No recommendation regarding suggestions and usage as per functional needs in the context.
- Actual context-awareness prototypes are different from the fundamental or theoretical concept of context representation in existence. This can be performed by analyzing the real-time healthcare projects.
- It is very difficult to build an efficient computerized system for healthcare as per the customer's perspective.

These days a mobile phone has become part and parcel for several routine tasks. Everybody depends on the mobile device for everything such as shopping, paying bills, healthcare, watching movies. Mobile phones or smartphones have become one of the significant means of communication in daily lives. The International Telecommunication Union (ITU) report depicts that the population in the world who are accessing cellular networks with mobile phones/smartphones are increasing monstrously. The dependency of people on mobile phones has become even more owing to enhanced embedded wireless communication techniques (GSM, Wi-Fi, Bluetooth, and others) for more reliable communication between two parties. Also, smartphones support a variety of new applications for providing more flexible

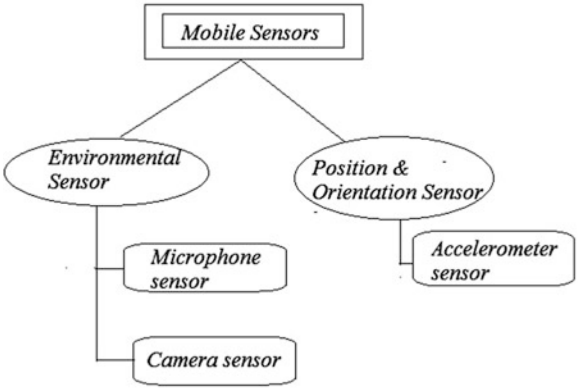


Fig. 4.2 Embedded sensors in modern mobile devices

options for users. With the help of the Internet, we can integrate healthcare services with mobile phones, to communicate healthcare centers on time. Practical medicine, online doctor consultation, and healthcare services all are supported by mobile devices, thereby providing mHealth solutions. Recently, mHealth filed is growing fast to create new applications for providing better healthcare services to all users. ThemHealth supported apps are available on webportals, such as Google Play Store. This makes its users to easily download and use the healthcare solutions efficiently on the mobile mode. Another most important factor is to determine the applicability of mobile phones in the healthcare area. In recent days, the upgraded mobile device is integrated with enhanced embedded sensing equipment, such as accelerometer, microphone for voice, and camera for imaging. These sensors intend to provide a better use of healthcare applications.

The basic applicability of a mobile sensing device is *mobility* with fewer handoffs or connection drops. Mobility is the most important factor nowadays so that users can track their health every time and anywhere. The integrated sensors in the mobile phone may be subdivided into two categories as in Fig. 4.2.

- Environment Sensors: These are used to perceive different characteristics of the mobile environment. Environment sensors include microphones, cameras, object trackers, etc.
- Position and Orientation Sensors: These sensors may include accelerometer, digital compass, gyroscope, and GPS. These electronic devices sense the location and orientation of the mobile phone with respect to time and space.

4.1.3 Applicability with Mobile Sensing Devices

Intelligent sensing devices offer dynamic extensions to static information-based healthcare systems with additional contexts. Context means instantaneous details associated with the surrounding environment, event, or entity, possessing specific attributes. The contextual attributes are referred to those characteristics that determine the situational cases or instances of a system or organization during a time frame. Contexts can be relatively static or dynamic concerning time. A context-aware system behaves intelligently to model user requirements adaptively and more realistically. Some contextual attributes include time, date geo-location, type of connectivity, user identity (name, password, designation, etc.), type of information, level of access, propose of visit/access, etc. Sources of such contextual attributes could be drawn from manual entries (drafted from a predefined list of options), sensors hardwired in computing devices (for tracking time, date, and location-related information), external sensors embedded in the environment, system state of execution, and other context provider services. The following sections illustrate some mobile sensing devices being widely used in health-related applications [7–10].

4.1.3.1 Microphone Sensor

Microphone sensors can be used in several ways in mHealth applications, apart from its inherent usage for communication and transmission. The application of the microphone involves:

- Providing communication and training platform for healthcare workers or professionals
- Diagnosing and imparting treatment support
- Tracking of diseases

This sensor offers an efficient way to use it in mHealth application so that a proper healthcare service will be delivered to customers.

4.1.3.2 Camera Sensor

The web camera or any motion detection sensor in the mobile phone is used to get useful information about a patient in form of images and videos that data will use in those applications where remote doctor consultation is needed.

4.1.3.3 Accelerometer and Geo-Location Facilitator Sensor

The main application in healthcare services is to track a user's physical activity level regularly. This application is very important for chronic disease patients because it reduces the risk of having any critical situation.

The embedded sensors and the mobility feature make the mobile device users in the field of healthcare service delivery. Context-awareness is gaining importance owing to the increasing opportunities being provided by cloud computing for online storage and instant availability of large-scale distributed data in homogeneous as well as in heterogeneous form.

4.1.4 Prevalent Applications

Some of the medical applications are enlisted that describes the significance of context-awareness in medical services:

4.1.4.1 Vocera

This was an initiative conducted in Birmingham, the USA by St. Vincent hospital for setting by an efficient system for communication [11]. This system is useful only for people using mobile devices. It is a wearable device for communication based on speech recognition. This system consists of an automatic calling facility, a small-size text screen that is flexible for enabling voice-call capabilities. The important features of this system are hands-free conversations and answering to perform calling, and storing voice message in case no answer to call is received. Biometric services are enabled device security with speaker identification and authentication. It transfers the relevant information to the users directly by using context-related information.

4.1.4.2 Mobile WARD

This is an innovation built by Aalborg University, Denmark, to maintain mobile electronic patient records. MobileWARD is the preliminary version of a device that is developed to sustain daily medical procedures in a hospital ward. Through this device, one can able to display the list of patients and their related information. The device can sense information and provide functionality basis on the location and time of hospital staff on duty. Patients will be chosen by a displaying list of patients or an active barcode at the hospital bed.

4.1.4.3 Other Medication Consumption Devices

Fishkin and Wang proposed a device for assisting medication facilities given at home [12]. It is a pad-like device that is specially designed for detecting:

- Lifting of a medicine bottle and putting it back
- Remaining number of pills from medicine bottles
- Postmedication suggestions
- Pills to take as per prescribed by the doctor or physician

4.1.4.4 Intelligent Hospital

Intelligent technology speeds up patient reports, diagnosis, and check-up summarization generation, by the intelligent device through context awareness [13, 14]. Some context-aware intelligent projects include:

- A hospital bed with context-awareness: With a built-in screen attached with the bed to display features as demanded by the patient (e.g., for viewing television) and also by doctors for accessing medical data of patients. The hospital bed has a sensor which could “sense” who is being allotted for the bed and what facilities are available in the nearby vicinity.
- A contextually aware container for pills intake: A self-aware container that contains pills and reports when it nears a particular patient. Moreover, containers could get labeled with the names of the patients and the respective dosage, the number of pills left, etc.
- A context-based electronic record for patients: The bed “sense” contexts, such as the patient, nurse, and the tray of medicine. It can also display other significant information as per the context, such as patient records, medicine schema.

4.1.4.5 Intelligent Wheelchair for Disabling Humans

It is an intelligent device with an accelerometer, camera, location sensors, and obstacle sensors. It allows the paralyzed person to move freely without the help of another person. It allows users to move independently. It contains an alarming system also which allows the users to call another person if they need it.

4.1.4.6 Dot Smartwatch

It is an intelligent wearable smartwatch that is easily affordable. This device helps blind people to access the messages, e-mails, tweets, books anywhere, and at any time.

4.1.4.7 MotionSavvyUNI

It is a two-way tool for communication by deaf people. It uses speech recognition technology and gesture technology for detecting the movement of hand and fingers with help of specialized camera and then it converts the signals into text in a while. In that way, it gives the meaning of sign language used by a deaf person. It has one more feature that is voice recognition. It converts speech into text for enabling both ways of communication. It further enables users to create their sign language and add custom dictionaries for custom sign language.

4.1.5 Open Challenges

Context-aware computing has some technical challenges that may occur during the implementation of healthcare service [15]:

4.1.5.1 Localization

Localization is the most important factor in mHealth solutions because it tracks the present location and provides relevant information depending upon the location of the user. For example, if the user lives in hilly areas where he takes fresh breathing air and suddenly arrives in the mainland area with too much pollution, he might not feel comfortable. mHelath solutions could prove beneficial to provide proper solutions according to his location what to do and how to survive in such an area.

4.1.5.2 Connectivity

Connectivity is the main concern also in the healthcare area. Most of the entire healthcare service is real-time so they need an Internet connection to sense the real-time data. Due to connection drop, one might fail to perceive real-time data and cannot provide proper healthcare service, promptly.

4.1.5.3 Real-Time Data

Real-time data is important through which service providers can provide accurate and proper service based on real-time conditions of users. For example, if blood pressure shoots up or steeps down from the normal level, mHealth solution should be able to react on the real-time input and respond with controlling and preventive measures. This may also involve informing family members or local emergency centers. Such attempts become infeasible in the absence of real-time data.

4.1.5.4 Environmental Issue

This issue often comes across when the device is based on the environmental condition. For instance, intelligent devices embedded in hospital beds that can sense information when some patient is assigned with all types of reports, including medicine schema, dietary chart, and so on.

4.1.5.5 Feasibility of Data

This means that the device can remark the feasibility of data. Data should be relevant and in the proper format so that doctors and service providers can provide a solution in less time. Data should not contain noise or redundancy.

4.1.5.6 Storage of Relevant Data and History

Such activity becomes very crucial in healthcare service, especially to perform a case study of patients and treat them more suitably. They can provide the exact solution or medicine after examining through the medical history of the patient and present real-time data.

4.2 Context Awareness in Healthcare

The context-aware intelligent healthcare framework provides the solution to one of the prevalent challenges confronted by traditional medical units of providing improved services to a large number of patients with constrained financial resources and manpower.

4.2.1 Smart Sensory Devices

Some of the important sensory devices used for smart healthcare delivery are enlisted as follows:

(a) Heart-Monitoring System: Smart sensors are installed unobtrusively in the patient's body for monitoring the heart. This sensor senses the heart rate, regular functioning of the heart. This system helps to prevent high death casualties due to cardiovascular disease. Based on the sensing information, the medical staff may provide treatment in advance so that the delay in treatment can be avoided if the patient is suddenly in a critical condition.
(b) Detection of Cancer: Today, cancer is one of the most critical threats to human life. The number of such patients is rising very fast every day. With a sensor capability to identify the nitric oxide that is released by the cancer cells, one can assist doctors to identify cancer cells at an early stage. These sensors are capable to distinguish the cancerous cells and based on collecting information the medical staff can detect the stage of cancer along with its possibility of treatment and cure.
(c) Hip-Guard: This system is designed for those patients who are improving from hip injuries. The embedded sensors can detect the patient's leg and position of

the hip along with the extent of rotation. Alarm signals are sent to the patients if any hip or leg position and rotation turn out to be wrong. This sensor provides real-time updates for better recovery.

(d) Asthma Sensors: Such electronic devices are used to sense agents in the environment that causes allergic reactions and report the present status constantly to the physician and patient so that necessary precautionary actions could be time taken.

(e) Mobihealth: Such an application is often deployed for continuously monitoring the patient outside the hospital environment. A sensor is used in this system to create a general platform for enabling healthcare in the domestic environment. Intelligent sensors detect the patient's conditions and assist accordingly to diagnose disease straightaway.

(f) Glucose Level Monitoring: Diabetes is a chronic disease that can lead to several chronic diseases, such as heart attack, stroke, high blood pressure, kidney disease. A diabetes sensor is capable to monitor the level of glucose and periodically send the results to a wirelessly operating digital assistant, and also provide one of the functionalities to inject insulin automatically whenever glucose threshold level is violated.

(g) E-Health Sensor Shield: In consists of different sensors, which include recording pulse, the oxygen level in the blood, breathing condition, temperature, galvanic skin response, blood pressure and electrocardiogram, and a glucometer.

(h) Encryption Algorithm Embedded Sensor: A sensor is used in this system to safeguard the confidentiality, integrity, and authentication of the collected data.

(i) 3G connectivity, Wi-Fi, and Bluetooth Sensors: These sensors allow sending the collected information at the same time. These sensors are very important for real-time data and to store it in a remote location.

(j) Environmental Monitoring Sensors: These devices are quite significant for analyzing the weather forecasting data and for quantifying the several environmental threat that can be averted before they occur. Environmental monitoring sensor collects data from a geographical region, which allows monitoring minute variations in the environmental parameters. Input from such sensors acts as crucial for suggesting mHealth solutions to patients.

4.2.2 Wireless Medical Sensors: Requirements and Challenges

The medical sensors must fulfill the following requirements, including wearability, interoperability, reliability, and security.

- Wearability: For employing inconspicuous and noninvasive monitoring of patients, the wireless medical sensors must be light in weight and small in size, so that it can easily wearable.
- Reliable Communication: Reliability of communication by medical sensors depends upon the need of sampling rates (from less than 1 to 1000 Hz). It can be

maintained by transferring the relevant information after applying data mining techniques on raw data.

- Security: It is an important requirement in medical sensors. To ensure the integral contents of patient-centric information gathered for medical examination. This further ensures the privacy of data to be maintained.
- Interoperability: The medical sensors operating wirelessly must allow users to easily construct a robust communication based on the state of health of the patients. Such information needs to be transferred reliably and rapidly to avoid the expiration of data freshness.

4.2.3 Context-Aware Sensor Data

Context-awareness describes an application that tends to make use of some context. Context-based data often helps to determine the context related to a person or object. The context-aware systems deploy mobile or wearable devices that are embedded with smart sensors particularly to monitor the current environmental situations to assist humans to maintain a suitable quality to their standard of living. The architecture of the context-aware system allows them to extract low-level context from a realistic heterogeneous physical world. Based on low-level context information acquired by sensors, high-level conceptual models are developed by the concerned authority. The following section provides a layout of key components and modeling processes in a context-aware system.

(a) Contextual Information: In the real world, sensing data and communicating mined and processed information is the key element of human interaction. The context can be defined as sensing the information and uses this information for providing better interaction between communicating parties in a real-world environment. Context-based information helps to reveal the joint-impact of all influencing contexts associated with a person or an object.

The context can be interpreted in different ways because it is completely dependent on the usage of smart sensors. The context is the assumption of a specific entity in the situation such as user profile, interaction, activity, user location. In that way, the concept of context is more powerful, more useful, more personalized, especially for mobile users.

(b) Context Representation: The context-awareness is applied to a mobile device to reduce human intervention by providing assistant services. The physical factors such as temperature, light, and location are also added in context-awareness systems, to perceive real-time contexts. The context represented in heterogeneous form or structure presents the collected information of the entity owing to situation-specific conditions of the real world. Following category of sensors are used to sample contextual data:

  - Physical sensors: Sensors in this category are capable of sensing physical environment data, for instance, location detection through GPS sensors, etc.

- Virtual sensors: Such sensors obtain data that is manually fed by users about applications/services via social networking portals.
- Logical sensors: Such types of sensors buffer extra information on physical and virtual sensors which is recorded from user interaction history and records.

  The context can be divided into:

- Device context: This includes Internet connectivity, cost of communication, and other resources, etc.
- User context: Such contexts include user profile, user geographic position or location, user activities in a particular situation, etc.
- Physical context: This captures temperature, noise, light intensity, traffic conditions, location, etc.
- Temporal context: Time factors such as day, week, month, and year are recorded as temporal context.

(c) Context Modeling: Context modeling aims to reduce the complexity and usability of the system/application and improve the maintainability and adaptability of the system or application for future use. This facilitates acquiring context-aware sensor data and thereby, model data according to the use.

4.2.4 Pervasive Healthcare

Pervasive healthcare systems provide telemedicine and healthcare services by caregivers and doctors. Patients can communicate with healthcare professionals at anywhere and anytime via the pervasive healthcare system. The main features of pervasive computing in the healthcare area include:

- Context-awareness
- System intelligence
- Knowledge-based discovery of database
- Privacy and security
- Mobility

Classification algorithms are used to extract the relevant data from a large data set. Such a colossal form of data undergoes through mining techniques, over which the decision-making process is executed. This further involves machine learning algorithms, real-time data collection, clustering algorithms for sensing intelligent locations, etc., for designing a better smart home healthcare system.

Pervasive computing has a mobility feature that permits users to move with their mobile devices such as Tab, Pad, cell phone, laptop and can access any service (G-mail, contacts, notifications, chat) via Internet/network at anytime and anywhere.

The objective of pervasive computing as highlighted in Fig. 4.3 helps to develop a refined and reliable service that is equipped with smart sensors embedded over

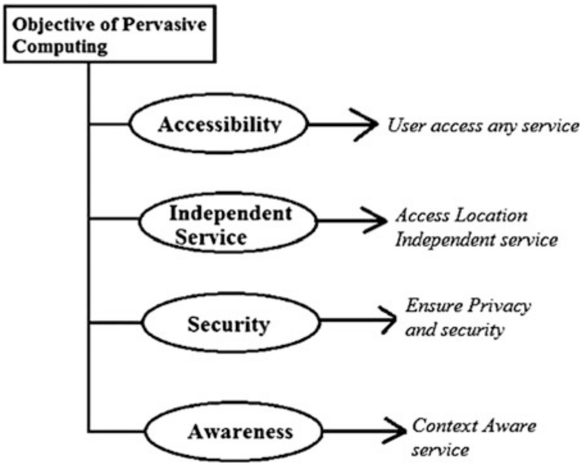


Fig. 4.3 The objective of pervasive computing

mobile devices. The pervasive healthcare systems can be classified into two main categories:

- Control and assisting for prevention systems
- Rescue in emergencies

It controls and assists the best measure for prevention. These guidelines are being provided for controlling the chronic disease and assisting with the best possible treatment for the prevention of patients from critical and adverse situations [16].

(a) Real-time Affordable Cardiovascular Emergency Detection System

This system enables handling patients suffering from chronic heart diseases. The cardiovascular emergency detection system is a wearable strap that can check the heartbeat rate timely and also draw conclusions from heart ECO. This device uses Bluetooth/Wi-Fi sensors for providing communication between healthcare providers and heart patients. This device can provide communication anywhere and at any time so that patients can be prevented from a critical situation. This system provides three types of warning messages with the help of smartphones. The system generates the following warning messages when required:

- Sending SMS alert which shows highlights the location of monitored patients
- Firing an alarm when the monitored patient's heartbeat rate exceeds the allowable optimal rate, that is, falling or rising beyond the acceptable threshold
- Sending alert messages to the emergency unit of the hospital to send an ambulance with medical facilities to rescue the monitored patient suffering in a critical state
- All such data and information are stored on the server. Data analysis application is applied to server data for decision-making purposes. This

server is directly connected to the emergency unit to promptly trigger the rescue process based on the received information.

(b) Monitoring system for diabetic's patients

The basic idea behind this system is to monitor diabetic patients. This system measures the sugar level in the blood of patients regularly. This system has mobile integration technology for sensing tasks. The patients can perform self-controlling and monitoring parameters by measuring the level of sugar in the blood using their mobile phones and in that way, they can reduce the risk of critical health conditions. This system provides many check-ups such as pulse rate, glucose level, BP, the weight of the patient and instantly provides measuring information using the smartphone. The capabilities of this system are:

- To send data report by the patient in real-time
- To transmit received data on a central server with the help of a smartphone
- To analyze stored server data by the expert team
- To make any decision (high level, very high level, diabetic or not) based on analyzed data

(c) Rescue operation in an emergency state

To achieve good health rescue systems that depend on good management of emergency services (EMS) and emergency vehicles (EMV). The management of emergency medical services follows ironic steps and takes actions to determine the priority for achieving a good precise decision. The EMV has route guidance and reception calling features. This assists the management and reduces the overall waiting time to rescue the patient, without much time elapsed.

(d) *Advanced system based on rescue operation in emergencies:* This system provides rescue operations for attending to sudden accidents. This application has four components:

- EMV: an emergency vehicle (like an ambulance)
- Traffic Management: to manage road traffic, that is, finding those areas which confront less or no traffic
- Geographic Information System: to give geographical information of accident area
- Conference Calling: to provide basic treatment in EMV through conference

This application is designed with the sole objective to reduce the EMV arrival time on accident areas, to reduce the required transfer time of patients from the accident spot to hospital, and to provide a high quality of healthcare service in EMV. This is one of the best instances of a real-time system for saving the life of people with a high success rate.

4.3 Mining Context-Based Health Aspects

Here, mining context-based health aspects means "knowledge discovery of the entire context that is related to health." Mining context-based health aspects refer to select the relevant data from the environment so that mined data will be used to determine the current health condition of users, that is, whether it is normal or critical!! For determining the health condition of the user, any application, or system needs a huge variety of data for analysis, data must contain all the aspects of health so that mining will be applied on all context data to get the relevant amount of information for generating the actual result.

4.3.1 Multiaspect Context-Based Dataset

Multiaspect context-based dataset refers to the sensing of multiple aspects of data (variety of parameters) from the current environment. Here, we define the multi-aspect of the context-based dataset by using some applications. An elaborated study is illustrated on these applications sense different parameters by context-awareness property and then maintain a context-based dataset for analysis purpose.

- To maintain blood pressure normal

  Nowadays, hypertension and stress are the most common chronic diseases all over the world. So many people are suffering from Blood pressure problems. Most of the patients use technology to monitor their BP regularly and for taking some remedies or precautions to get rid of high BP. Withings is one of the connected devices, which is easily available on the market. An app consisted device which is connected to Bluetooth for monitoring the blood pressure itself. This app measures all the relevant aspects of health which will be helpful to monitor the best result. This app measures not only blood pressure but also check the heartbeat rate and also counts the footsteps the person takes weekly. The app takes these three measurements and reports based on collecting health aspects with the medical recommendations. This app has a useful alarming feature so that patients can take his/her medicine on time and patients can set an alarm for monitoring BP. In that way, this app sensed multiaspects context data set and reports the actual result after mining.
- To keep healthy heart

  We know our heartbeats 100,000 times per day. Many gadgets are in the market to keep track of the heartbeat rate and analyze the health of the heart. In that way, users of the gadget can predict and prevent lots of heart problems such as sudden heart attacks. AliveCor Heart Monitor is an electrocardiogram with the help of a mobile phone. It is easily attached to the phone case. This device is slim and tiny in size. It records ECG, past the patient's data. It takes all aspects of health parameters to check whether the patient's heart is working normally or not! This device takes the history of the patient to report the best result on time. This app stores all aspects of health as a context dataset. Later this dataset will be mined to get the relevant information for reporting good health.

- To measure body temperature:

  To maintain the optimal body temperature is very important. Through thermometer measuring of body temperature is very frustrating and takes long minutes of sitting. Recent innovations allow us to measure the temperature and also other parameters in less time. Viatom Checkmethe world's first medical Tricorder which diagnoses medical conditions within seconds. It is a proper medical multitool which not only measures body temperature but also records ECG, pulse rate and oxygen saturation, BP, physical activity, and sleep. This device is used to check the body temperature and all other aspects of health and provide the result whether the patient should have to visit a physician or not!! On the basis of the report, the patient can plan his day.

The above applications help us to understand how a multiaspect context-based dataset is created and how this dataset to help healthcare providers to generate the result. The huge collection of datasets may contain some noise and irrelevant data that will not be useful. For extracting the relevant and useful data from the large dataset, we apply mining techniques like clustering.

4.3.2 Collaborative Intelligent Mining

Intelligent mining or data mining refers to extracting or mining knowledge (relevant information) from a huge amount of data. It is a logical process to search for a pattern from the large dataset. The main goal of mining is to find the pattern which will be used in the future to make the right decision. In other words, intelligent mining is a knowledge discovery of data.

Intelligent mining is an iterative process. In Fig. 4.4, we define the overall procedure of collaborative intelligent mining [17, 18]. The following steps of this process highlight working in brief.

- Data wiping: It is the process to remove noise and inconsistent data from the large dataset.
- Data amalgamation: It is the process to combine multiple data sources.
- Data selection: It refers to analysis of relevant data that are retrieved from the database.
- Data alteration: Selected data are altered into other appropriate forms by applying mining rules.
- Data mining: It is a logical process to search a pattern from the large dataset.
- Pattern evaluation: It identifies the best pattern for the best prediction.
- Knowledge presentation: It represents and visualizes the knowledge or pattern for the user by applying some techniques.

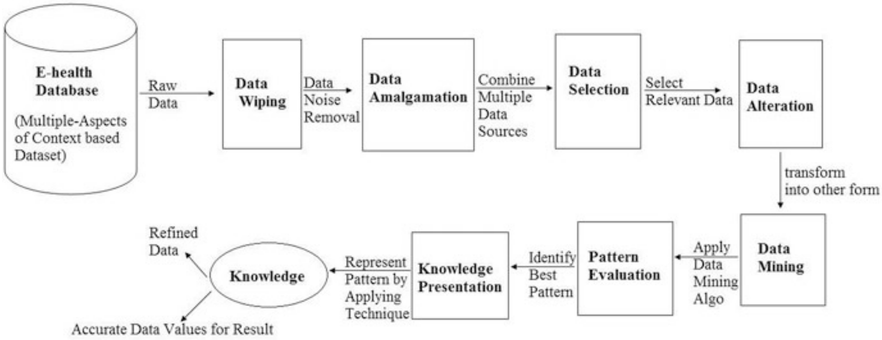


Fig. 4.4 Collaborative intelligent mining

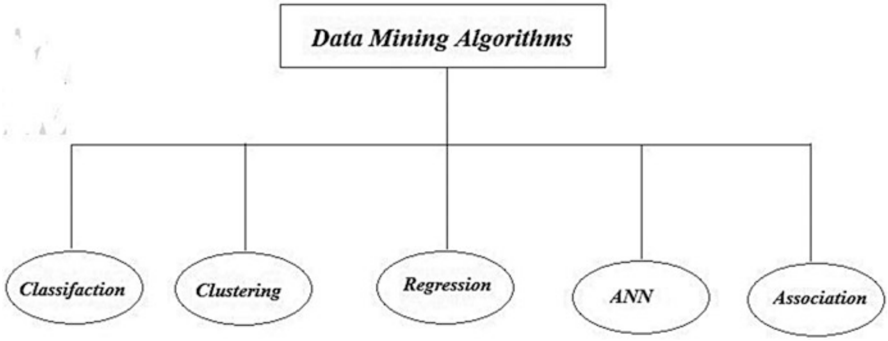


Fig. 4.5 Popular data mining algorithms for mHealth applications

Many different types of algorithms and techniques are available in intelligent mining for mining the relevant information for the user from the large dataset. Data mining algorithms and techniques are Classification, Clustering, Regression, Artificial Neural Networks (ANN), Association Rules, K-means, etc. Some Popular Data Mining Algorithms for mHealth Applications are shown in Fig. 4.5.

(a) Classification

Classification is the most commonly used intelligent data mining. Predefined examples are stored to develop a classifier that can classify the large recorded dataset. This technique can be used with neural networks and decision trees also. It is the process that involves learning by example and then classification. This technique is based on two phases:

- Learning phase: Classifier uses the training data set and then analyzed the data by classification rules.
- Classification of data: The tested data can be used to estimate the accurate classification rule for finding the best pattern or relevant data.

A classifier is just a model to perform the above two phases and provide the required and relevant pattern after applying accurate classification rules. In other words, we can say that it classifies the knowledge based on the features in the large dataset. Types of classification models:

– Decision tree classification
– Bayesian classifier
– Artificial neural intelligence
– Support vector machines
– Classification with associations rules

(b) Clustering

It is the process to identify the homogeneous classes of objects. The clustering technique helps to identify dense and spare regions of object space. It can find the distributive pattern and co-relates that pattern with the dataset. The classification works on heterogeneous data but clustering makes classes of homogeneous data. Clustering makes a class of those data objects which has similar functionality [19]. The clustered data becomes the refined knowledge discovery of large data. The different types of clustering methods include:

- Partitioning techniques
- Hierarchical clustering
- Density-based clustering methods
- Grid computing

(c) Regression

The regression technique is based on predication. This technique is used to define the relationship between the dependent and one or more independent variables. Independent variables are known attributes and they help to predict the next variables. It analyzes the relationships among the variables. It models a set of dependent data and independent data from the large dataset which is further useful for maintaining the best result. Types of regression methods are:

- Linear and logistic regression
- Multivariate linear regression
- Multivariate nonlinear regression

(d) Association rule

Association and correlation are used to find frequent item set among large data sets. This frequent item set helps to make better decisions. Association rule algorithms help to generate new rules with values less than 1. We can generate so many numbers of association rules for a given dataset with certain values. Types of association rules include multilevel, multidimensional, and quantitative.

(e) Neural networks

A neural network is a set of interconnected input and output units (called neurons). Each interconnected units has weight and activation function. Each neuron has training data to fire on particular input patterns. Neuron works on two modes: learning and training mode, using mode. In the learning phase, networks adjust the neuron weights so that the correct result can be predicted. Types of neural networks broadly involve:

- Backpropagation
- Single-layer perceptron
- Multilayer perceptron

4.4 Security in Healthcare Service Delivery Model: A Case Study

Context mainly relates to the situation concerning certain influencing factors. Contextual attributes are those characteristics or attributes that determine such situational cases of a system or organization. This section highlights a case study on context-aware security in healthcare services that aim to safeguard resources based on the present context of an entity (person, object, profile, or application) (Fig. 4.6). Context-aware security is gaining significance owing to widespread research in distributed data, cloud computing, sophisticated crime, cyber threats.

Different forms of security could be triggered for access control, authentication, authorization, encryption, etc., of records and information. Traditional security is coarse and context insensitive, and relies on consistent configuration (Fig. 4.7). Modes of context-awareness could be decided based on static or dynamic contexts. Depending upon which contextual cluster on entity belongs will decide the type or level of security is being imposed. Contextual attributes include time, date, geolocation, type of connectivity, user identity (designation, user id, password), type of information, level of access, purpose. Source of contextual attributes could be from manual entries (user comment or from a predefined list of options), sensors embedded in a computing device (for recording instantaneous information, like time, date, location), extended sensors in environment, system state, other context providing service, etc.

4.5 Recent Trends in Intelligent Healthcare Delivery

Recent trends include some intelligent healthcare services are cloud-based which stores data on the cloud and IoT in real-time. Here, we are explaining some intelligent healthcare services; these are the following [20]:

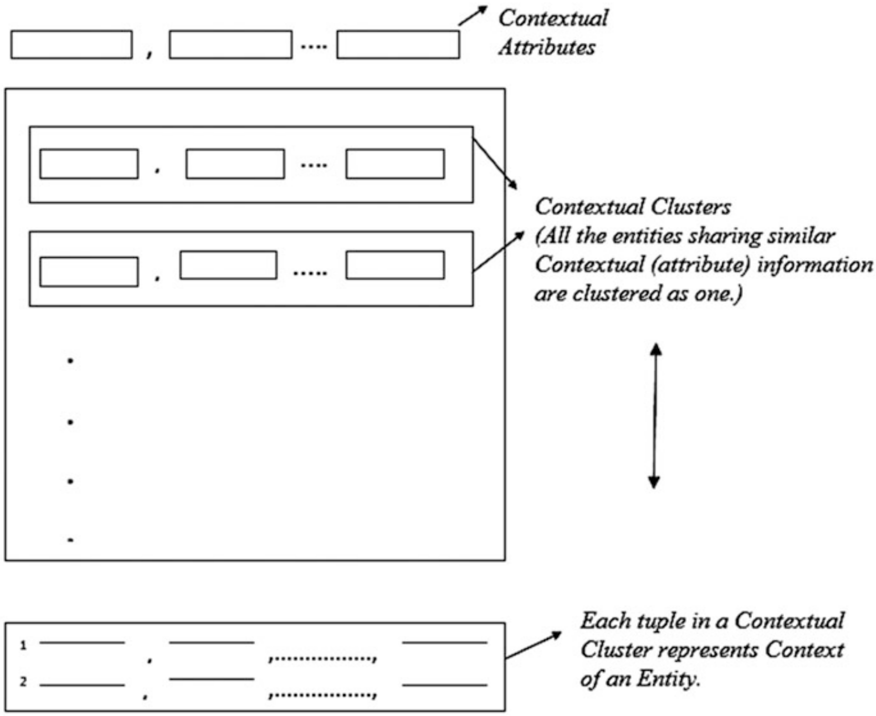


Fig. 4.6 Instances of contextual attributes and clusters for context-aware security

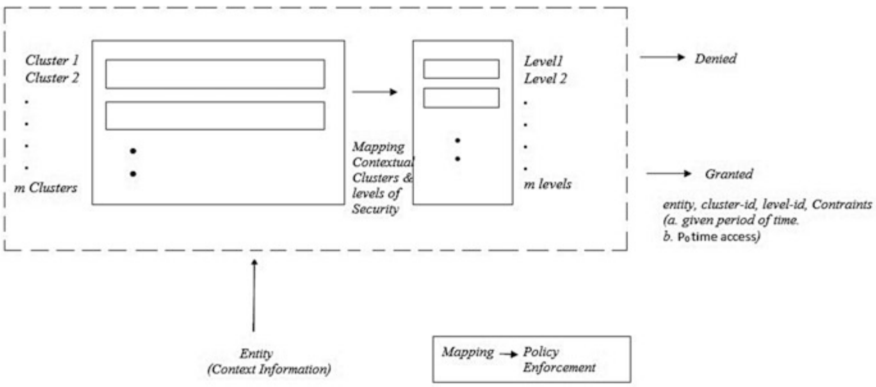


Fig. 4.7 Instances of mapping between contextual clusters and security levels

- *Intelligent wheelchair for disabling humans:* It is an intelligent device with an accelerometer, camera, location sensors, and obstacle sensors. The device allows the paralyzed person to move freely without the help of another person. It allows users to move independently. It contains an alarming system also which allows the users to call another person if they need it.

- *Dot: It* is an intelligent wearable smartwatch that is easily affordable. This device helps blind people to access the messages, e-mails, tweets, books anywhere, and at any time.
- *UNI:* UNI is a two-way communication tool for the deaf. It uses speech technology and gesture technology for detecting the movement of hands and fingers with help of specialized camera and then it converts the signals into text in a while. In that way, it gives the meaning of sign language used by a deaf person. It has one more feature that is voice recognition. It converts speech into text for two-way communication. It also enables users to create their sign language and add custom dictionaries for custom sign language.
- *Cancer Detection:* Today's cancer is the biggest threat to human life. The number of cancer patients is rising very fast day by day. A sensor can detect the nitric oxide which is emitted by cancer cells. This sensor can be placed by doctors where they identify the cancer cells. These sensors are capable to differentiate cancerous cells between different types of cells and based on collecting information the medical staff can detect the stage of cancer at an early age.
- *Fitbit Aria:* A smart gadget to keep a track of body fat and your weight. This is a cloud-based gadget which stores measuring data on cloud server through the Wi-Fi network. It creates a separate account for the user on the cloud and maintains his/her record separately and senses the historic data also to provide the best result. With the help of this app, the user can compare their changes in weight by doing physical activity.

4.6 Conclusion

Currently, there are several prevalent healthcare applications and devices that are connected with a mobile platform to monitor human health. These are intelligent applications that work in real-time and offer dynamic extensions to the static healthcare frameworks with contextual processing of relevant medical data. Context refers to instantaneous and scenario-specific details associated with the surrounding environment or entity that enables to pour extensive insights on the cause and remedy of an ailment. Our work introduces the utility of such context-aware systems, specifically in terms of real-time adaptability.

To summarize, our research addresses the principal methods, models, current trends, and further projections of future research in intelligent healthcare. This would foster the understanding of formidable challenges confronted in developing protocols and standards. Our research would serve as a leading reference for those who seek broad knowledge with in-depth conceptualization of context-awareness in smart healthcare delivery.

References

1. N. Bricon-Souf, C.R. Newman, Context awareness in health care: A review. Int. J. Med. Inform. **76**(1), 2–12 (2007)
2. L. Catarinucci, D. De Donno, L. Mainetti, L. Palano, L. Patrono, M.L. Stefanizzi, L. Tarricone, An IoT-aware architecture for smart healthcare systems. IEEE Internet Things J. **2**(6), 515–526 (2015)
3. S.R. Islam, D. Kwak, M.H. Kabir, M. Hossain, K.S. Kwak, The internet of things for health care: A comprehensive survey. IEEE Access **3**, 678–708 (2015)
4. G. Chen, D. Kotz, A survey of context-aware mobile computing research. Dartmouth Comput. Sci. Tech. Rep., TR2000–TR2381 (2000)
5. M. Wang. A Flexible, Low-Overhead Ubiquitous System for Medication Monitoring. Intel Corporation, Tech. Rep (2003)
6. M. Satyanarayanan, Pervasive computing: Vision and challenges. IEEE Pers. Commun. **8**(4), 10–17 (2001)
7. S.C. Mukhopadhyay, Wearable sensors for human activity monitoring: A review. IEEE Sensors J. **15**(3), 1321–1330 (2014)
8. T. Islam, S.C. Mukhopadhyay, N.K. Suryadevara, Smart sensors and internet of things: A postgraduate paper. IEEE Sensors J. **17**(3), 577–584 (2016)
9. M. Chaudhari, S. Dharavath, Study of smart sensors and their applications. Int. J. Adv. Res. Comp. Commun. Eng. **3**(1), 5031–5034 (2014)
10. N.B. Roy, D. Das, Role of body area sensor networks in smart health care. Int. J. Future Comput. Commun. **4**(5), 320 (2015)
11. V. Stanford, Beam me up, doctor McCoy [pervasive computing]. IEEE Pervasive Comput. **2**(3), 13–18 (2003)
12. M. Wang. *A flexible, low-overhead ubiquitous system for medication monitoring*. Intel Corporation, Tech. Rep (2003)
13. J.E. Bardram, Hospitals of the future–ubiquitous computing support for medical work in hospitals, in *Proceedings of UbiHealth*, vol. 3, (2003, October)
14. J.E. Bardram, Applications of context-aware computing in hospital work: Examples and design principles, in *Proceedings of the 2004 ACM symposium on Applied computing*, (2004, March), pp. 1574–1579
15. M.J. Covington, W. Long, S. Srinivasan, A.K. Dev, M. Ahamad, G.D. Abowd, Securing context-aware applications using environment roles, in *Proceedings of the sixth ACM symposium on Access control models and technologies*, (2001, May), pp. 10–20
16. H. Faiez, J. Akaichi. A review of pervasive healthcare from prevention to emergency rescue: Systems, tools, platforms and techniques. J. Med. Bioeng. **6**(1) (2017)
17. X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, et al., Top 10 algorithms in data mining. Knowl. Inf. Syst. **14**(1), 1–37 (2008)
18. M. Gera, S. Goel. Data mining-techniques, methods and algorithms: A review on tools and their validity. Int. J. Comput. Appl. **113**(18) (2015)
19. R. Xu, D. Wunsch, Survey of clustering algorithms. IEEE Trans. Neural Netw. **16**(3), 645–678 (2005)
20. P. Sarkar, D. Sinha. Application on pervasive computing in healthcare–a review. Indian J. Sci. Technol. **10**(3) (2017)

Chapter 5
Optimization of Training Data Set Based on Linear Systematic Sampling to Solve the Inverse Kinematics of 6 DOF Robotic Arm with Artificial Neural Networks

Ma. del Rosario Martínez-Blanco, Teodoro Ibarra-Pérez, Fernando Olivera-Domingo, and José Manuel Ortiz-Rodríguez

5.1 Introduction

The large amount of data available today constitutes one of the most valuable capital for organizations, because through its analysis, it is possible to obtain strategic information for decision-making, detection of behaviors, establishment of predictive models, among others [1].

The current volume of information has exceeded the processing capabilities of current conventional systems. The intervention of new processing algorithms and scalable techniques, which allow fast and efficient information processing, is increasingly necessary [2].

When having a very large amount of data, there are usually two ways to approach the problem. One way can be to redesign the algorithms without affecting performance to obtain an efficient execution with all the data. A second approach may involve reducing the data set to obtain a very similar result as if the entire volume of data were used [3].

Success in knowledge extraction algorithms is highly dependent on the integrity and consistency of the extracted data. Particularly in the field of artificial neural networks, most research focuses its efforts on specific applications and training algorithms that improve precision and convergence of results. However, most of the studies do not describe the procedure that was applied in the data preprocessing

M. d. R. Martínez-Blanco · T. Ibarra-Pérez · J. M. Ortiz-Rodríguez (✉)
Laboratorio de Innovación y Desarrollo Tecnológico en Inteligencia Artificial (LIDTIA), Universidad Autónoma de Zacatecas (UAZ), Zacatecas, México

T. Ibarra-Pérez · F. Olivera-Domingo
Unidad Profesional Interdisciplinaria de Ingeniería Campus Zacatecas (UPIIZ), Instituto Politécnico Nacional (IPN), Zacatecas, México

S. Paul et al. (eds.), *Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-77558-2_5

stage and, in general, the determination of success or failure in the knowledge extraction algorithms is influenced by the quality of the training data [4].

Data preprocessing is a stage that can increase the quality and reliability of the data because a low quality in the data leads to a low quality in the knowledge extracted. Although it is true that a very complete training data set would allow a better understanding of the problem to be obtained, the training time required will be much longer and computationally costly, making it infeasible [5].

Sampling is one of the most appropriate preprocessing methods to solve this problem due to its advantages in performance and low processing cost required during its application in knowledge extraction algorithms in various fields of engineering, statistics, machine learning, and data mining [6].

Data preprocessing techniques focus on two areas: data preparation and data reduction. Data preparation is mandatory and refers to the adequacy of the data so that the algorithms can be executed correctly, such as normalization, cleaning, and probably the recovery of lost data. On the other hand, data reduction is not always required and refers to the generation of a reduced size that maintains the integrity of the information as much as possible, such as the selection of characteristics, the selection of instances, the grouping, and sampling among others [7, 8].

Linear Systematic Sampling (LSS) is one of the most used techniques thanks to its ease of use. It was first introduced in 1944 by [9] and is also known as Cluster Sampling method, which consists of dividing the population N into k groups of n elements each, allowing all the elements of the data set to have the same chances of being selected [10, 11].

In this study, a Reduction Data Filter (RDF) algorithm based on the LSS method was implemented to process more than 4 billion data in a conventional AMD Ryzen 7 processor with 16 GB of RAM. In this way, a training data set was generated, which was used by two different Artificial Neural Networks (ANN) architectures to analyze the performance and generalizability of both architectures.

5.1.1 Artificial Neural Networks

The understanding of our brain has allowed the creation of Artificial Neural Networks (ANN) due to the fact that they are inspired by its operation, while a microprocessor processes information sequentially, the human brain does it in parallel and concurrently, however, we are still far from emulating the simplest capabilities of human reasoning, although ANN is currently a very powerful instrument for various applications in engineering and a very promising field of research [12].

An artificial neuron can be described as a computational structure or a mathematical abstraction model inspired by the neurons of our brain to which input signals arrive, as occurs with the dendrites of a real neuron, and generates an output signal, as happens with the axon [13]. An example of the schematic of an artificial neuron is shown below in Fig. 5.1.

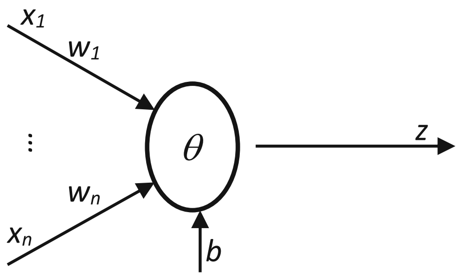


Fig. 5.1 Scheme of an artificial neuron

Each of the inputs $x_{1,...,}$ x_n is assigned a value or synaptic weight $w_{1,...,}w_n$. The output of the perceptron is an independent function of the value of its inputs, with their respective weights and threshold value. The weights w_i represent the intensity of synapses connecting two neurons and can be negative or positive. θ is the transfer function or firing threshold from which the neuron is activated [14].

The output is calculated by the cumulative sum of the product of all the input signals multiplied by their corresponding synaptic weights plus *a* bias, *b*, as shown below in Eq. (5.1).

$$z = \sum_i x_i \times w_i + b \tag{5.1}$$

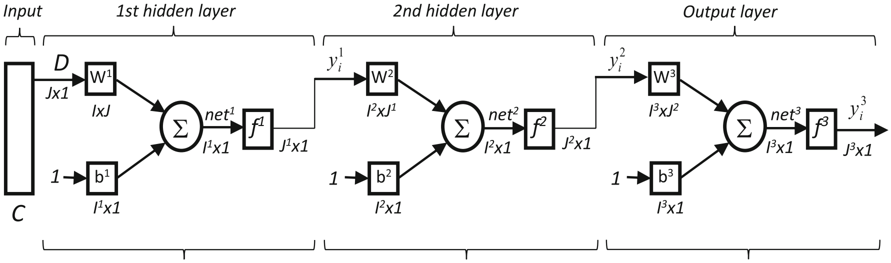
The output generated is used as input to a transfer function where the response to the artificial neuron is generated, as shown below in Eq. (5.2).

$$y_i = f\left(\sum_i x_i \times w_i + b\right). \tag{5.2}$$

ANNs are interconnected neurons distributed in parallel. Generally, the form of connection between neurons determines the type of network and they are usually grouped in layers. A layer is a set of neurons and according to its location it can be the input layer, a hidden layer, or the output layer. The architecture of a neural network is determined according to the arrangement of neurons within the layers and the forms of connection between them [14, 15].

Figure 5.2 shows a simplified model that describes an ANN with three layers, where it is observed that the network has R^1 inputs, i^1 neurons in the first layer, i^2 neurons in the second layer, and i^3 neurons in the third layer. The bias constant with a value of 1 is added in each neuron and in this way a layered approach can be obtained to analyze the complete structure of the network [16].

According to the feedback existing in the network, it can be seen that, if there is no connection between the output layer and the neurons of the input layer, the network does not maintain a previous memory state, so it is a forward propagation network. On the other hand, when there is feedback between the input and output



$$y_i^1 = f^1\left(\sum_{i=1}^{n} d_j w_{ij}^1 + b^1\right) \qquad y_i^2 = f^2\left(\sum_{i=1}^{n} y_i^1 w_{ij}^2 + b^2\right) \qquad y_i^3 = f^3\left(\sum_{i=1}^{n} y_i^2 w_{ij}^3 + b^3\right)$$

Fig. 5.2 Artificial Neural Network with two hidden layers

layers, a memory of their previous states is kept and the next state depends not only on the input signals but also on the previous states of the network, generating a backward propagation network [14].

In 1986 Rumelhart and McCelland proposed the Back Propagation Neural Network (BPNN) whose fundamental architecture consists of three layers. In this type of network, the number of layers that are necessary can be used and its mathematical foundation is based on the gradient descent algorithm, where the synaptic weights constantly change when processed by neurons through an activation function, producing outputs for the next layer until the minimum error is reached [14, 17, 18].

There is generally a problem with this type of network because the structural parameters must be established at the beginning of the training. Currently there is no procedure that guarantees the optimal configuration of these parameters and they are usually proposed according to the researcher's previous experience in trial and error experiments [19].

In Generalized Regression Neural Network (GRNN) architectures, unlike BPNN, it is not necessary to optimize the training parameters such as the learning rate and the momentum, on the contrary, the smoothing factor or propagation value of the network is determined. This value must be greater than zero and can generally be in the range of 0.01 to 1, obtaining excellent results. To determine the propagation value, it is necessary to estimate the probability density function according to the samples used in the training of the network and to carry out several experiments in order to determine the most appropriate value to train the GRNN [20, 21].

Donald F. Spech introduced this type of network in 1991, which is capable of producing continuous output values and rapidly draining sparse data sets. Due to the fact that only one forward propagation is necessary, the training of this type of network is very fast compared to the BPNN, where the information must be backward and forward propagated several times until an acceptable value of the desired error is found [20, 22].

5.1.2 Inverse Kinematics Solution with Artificial Neural Networks

Inverse kinematics is one of the most interesting problems in the field of industrial robotics. Inverse kinematics consists of determining the joint values of the end effector for a certain position and orientation in a Cartesian plane. This problem can be solved by means of a closed solution, since on some occasions the solution in real time is necessary for applications such as tracking a trajectory. However, the solution could be very complex due to the geometry of the manipulator, forcing the implementation of iterative solutions, which would be unfeasible for real-time applications, requiring the intervention of appropriate predictive models in the field of soft computing, where ANNs are one of those techniques that can be used to obtain acceptable results [23–25].

During the last decades, researchers have shown a special interest in the use of ANN, particularly due to its characteristics related to nonlinearity, parallelism, high robustness, fault tolerance, and great capacity for learning and generalization from complex and nonlinear examples [26–29].

The multilayer perceptron trained with the back propagation algorithm (BP) is one of the most widely used network architectures in modeling, optimization, and classification applications [19, 30–32]. This type of network is one of the most frequently used to learn the equations of the direct and inverse kinematics of 6 Degrees of Freedom (DOF) robotic manipulators, where the network learns the functional relationship between the input space and the output space through supervised training, based on a mapping adapting the solution in a nonlinear relationship between the locations of the end effector with the desired location [23, 33].

Currently, the determination of the structural parameters in the use of ANN continues to be a difficult and complicated task, because the parameters are generally defined by the researcher's previous experience in trial and error procedures, investing large amounts of time and resources, without guaranteeing the optimal configuration of the structural parameters [19, 28, 34, 35].

In this study, the Robust Design Artificial Neural Networks (RDANN) methodology was used to determine the optimal parameters in the first proposed network architecture: BPNN. Likewise, a systematic procedure was used to calculate the optimum value in the second proposed network architecture: GRNN. Both networks were used to solve the Inverse Kinematics (IK) of a 6 DOF robotic manipulator after training them with two data sets of the same size, but with different preprocessing characteristics, with the aim of analyzing the performance and generalizability of the proposed networks based on the quality of the training data [19, 36].

5.2 Neural Networks Based Inverse Kinematics Solution

5.2.1 Kinematics Analysis of Ketzal Robot

The morphology of robotic manipulators generally refers to the shape of the components and their structural mechanical parts [37]. A robotic manipulator generally has rigid mechanisms known as links that are connected by prismatic or rotating joints forming an open chain that can be operated by actuators [38].

In this study, the structure of a robotic manipulator called Ketzal with six DOF was used. Figure 5.3 shows the structure and coordinates reference systems of the robot, which is based on an open source, low-cost, and 3D printed project [39].

The forward kinematics calculation is about finding the position and the orientation vectors of the end effector with respect to a fixed coordinate reference system, given the vector of joint angles [40]. The inverse kinematics problem is about calculating the vector of joint angles given the position and orientation vectors of the end effector with respect to a fixed coordinate reference system [41].

The forward kinematics calculation results in an homogeneous transformation matrix T of size $4x4$, as shown in Eq. (5.3), where the spatial configuration between the joints of the manipulator is related to the position and orientation with respect to a fixed reference system [24].

$$T_0^6 = \begin{bmatrix} R_0^6 & P_0^6 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.3}$$

where R_0^6 represents a 3×3 size rotation matrix composed of the vectors n, o y a and that defines the orientation of the end effector and P_0^6 is the position vector of the end effector in the coordinate reference system. In this chapter, the Denavit-Hartenberg (D-H) method was used to calculate the forward kinematics of the Ketzal manipulator by implementing our basic transformations [42]. The D-H parameters are shown below in Table 5.1.

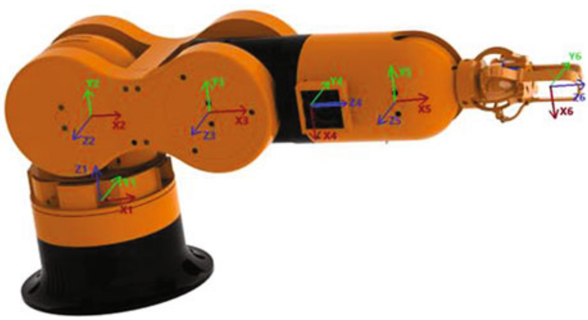


Fig. 5.3 Ketzal robotic manipulator

Table 5.1 Ketzal robot D-H parameters

Link offset (cm)	Joint angle (rad)	Link length (cm)	Twist angle (rad)
$d_1 = 20.2$	$\theta_1 = q_1$	$a_1 = 0$	$\alpha_1 = \frac{\pi}{2}$
$d_2 = 0$	$\theta_2 = q_2$	$a_2 = 16$	$\alpha_2 = 0$
$d_3 = 0$	$\theta_3 = q_3 + \frac{\pi}{2}$	$a_3 = 0$	$\alpha_3 = \frac{\pi}{2}$
$d_4 = 19.5$	$\theta_4 = q_4$		

The transformations, which depend on the configurations of the links, consist of a succession of rotations and translations allowing to relate the reference system of element *i* with the system of element *i-1*, which is given by Eq. (5.4) [43].

$$^{i-1}A_i = \mathrm{Rot}_{z\theta i}\mathrm{Trans}_{x,di}\mathrm{Trans}_{x,\alpha i}\mathrm{Rot}_{z,\alpha i} \tag{5.4}$$

Similarly,

$$^{i-1}A_i = \begin{bmatrix} C\theta_i & -S\theta_i C\alpha_i & S\theta_i S\alpha_i & a_i C\theta_i \\ S\theta_i & C\theta_i C\alpha_i & -C\theta_i S\alpha_i & a_i S\theta_i \\ 0 & S\alpha_i & C\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.5}$$

where i is the number of the link, θ_i is the angle of rotation of the joint, α_i is the rotation of the joint, a_i is the length of the link, d_i is the displacement of the link, and $C\theta_i = \cos(\theta_i)\, y\, S\theta_i = \sin(\theta_i)$.

Realizing the product of the six matrices obtained from Eq. (5.6), the matrix that indicates the location of the final system with respect to a reference system located at the base of the robot is obtained, which is known as the homogeneous transformation matrix T_0^6 [43].

$$T_6^0 = {}^0A_1 \cdot {}^1A_2 \cdot {}^2A_3 \cdot {}^3A_4 \cdot {}^4A_5 \cdot {}^5A_6 \tag{5.6}$$

where

$$^0A_1 = \begin{bmatrix} \cos(q_1) & 0 & \sin(q_1) & 0 \\ \sin(q_1) & 0 & -\cos(q_1) & 0 \\ 0 & 1 & 0 & a_0 + a_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.7}$$

$$^1A_2 = \begin{bmatrix} \cos(q_2) & -\sin(q_2) & 0 & a_2 * \cos(q_2) \\ \sin(q_2) & \cos(q_2) & 0 & a_2 * \sin(q_2) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.8}$$

$$^{2}A_{3} = \begin{bmatrix} -\sin(q_3) & 0 & \cos(q_3) & 0 \\ \cos(q_3) & 0 & \sin(q_3) & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.9}$$

$$^{3}A_{4} = \begin{bmatrix} \cos(q_4) & 0 & -\sin(q_4) & 0 \\ \sin(q_4) & 0 & \cos(q_4) & 0 \\ 0 & -1 & 0 & a_3 + a_4 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.10}$$

$$^{4}A_{5} = \begin{bmatrix} \cos(q_5) & 0 & \sin(q_5) & 0 \\ \sin(q_5) & 0 & -\cos(q_5) & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.11}$$

$$^{5}A_{6} = \begin{bmatrix} \cos(q_6) & -\sin(q_6) & 0 & 0 \\ \sin(q_6) & \cos(q_6) & 0 & 0 \\ 0 & 0 & 1 & a_5 + a_6 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.12}$$

Getting,

$$n_x = \sin(q_1)\sin(q_6)\cos(q_4) + \sin(q_1)\sin(q_4)\cos(q_5)\cos(q_6) + \sin(q_2+q_3) \sin(q_4)\sin(q_6)\cos(q_1) - \sin(q_5)\cos(q_1)\cos(q_2+q_3)\cos(q_6) - \sin(q_2+q_3) \cos(q_1)\cos(q_4)\cos(q_5)\cos(q_6) \tag{5.13}$$

$$n_y = -\sin(q_6)\cos(q_1)\cos(q_4) - \sin(q_4)\cos(q_1)\cos(q_5)\cos(q_6) + \sin(q_1) \sin(q_2+q_3)\sin(q_4)\sin(q_6) - \sin(q_1)\sin(q_5)\cos(q_2+q_3)\cos(q_6) - \sin(q_1) \sin(q_2+q_3)\cos(q_4)\cos(q_5)\cos(q_6) \tag{5.14}$$

$$n_z = -\sin(q_2+q_3)\sin(q_5)\cos(q_6) + \cos(q_2+q_3)\cos(q_4)\cos(q_5)\cos(q_6) - \sin(q_4)\sin(q_6)\cos(q_2+q_3) \tag{5.15}$$

$$o_x = \sin(q_1)\cos(q_4)\cos(q_6) + \sin(q_2+q_3)\sin(q_4+q_6)\cos(q_1) - \sin(q_1)\\ \sin(q_4)\sin(q_6)\cos(q_5) - \sin(q_5)\sin(q_6)\cos(q_1)\cos(q_2+q_3) \tag{5.16}$$

$$o_y = \sin(q_4)\sin(q_6)\cos(q_1)\cos(q_5) + \sin(q_1)\sin(q_2+q_3)\sin(q_4)\cos(q_6)\\ + \sin(q_1)\sin(q_5)\sin(q_6)\cos(q_2+q_3) - \sin(q_1)\sin(q_2+q_3)\sin(q_6)\cos(q_4)\\ \cos(q_5) - \cos(q_1)\cos(q_4)\cos(q_6) \tag{5.17}$$

$$o_z = \sin(q_2+q_3)\sin(q_5)\sin(q_6) - \sin(q_6)\cos(q_2+q_3)\cos(q_4)\cos(q_5)\\ - \sin(q_4)\cos(q_2+q_3)\cos(q_6) \tag{5.18}$$

$$a_x = \sin(q_1)\sin(q_4)\sin(q_5) - \sin(q_2+q_3)\sin(q_5)\cos(q_1)\cos(q_4) + \cos(q_1)\\ \cos(q_2+q_3)\cos(q_5) \tag{5.19}$$

$$a_y = -\sin(q_4)\sin(q_5)\cos(q_1) - \sin(q_1)\sin(q_2+q_3)\sin(q_5)\cos(q_4)\\ + \sin(q_1)\cos(q_2+q_3)\cos(q_5) \tag{5.20}$$

$$a_z = \sin(q_2+q_3)\cos(q_5) + \sin(q_5)\cos(q_2+q_3)\cos(q_4) \tag{5.21}$$

$$p_x = \Big[\sin(q_1)\sin(q_4)\sin(q_5) - \sin(q_2+q_3)\sin(q_5)\cos(q_1)\cos(q_4) + \cos(q_1)\\ \cos(q_2+q_3)\cos(q_5)\Big](a_5+a_6) + \cos(q_1)\cos(q_2+q_3)(a_3+a_4) + \cos(q_1)\\ \cos(q_2)\,a_2 \tag{5.22}$$

$$p_y = -\Big[\sin(q_4)\sin(q_5)\cos(q_1) - \sin(q_1)\cos(q_2+q_3)\cos(q_5) + \sin(q_1)\sin(q_2+q_3)\sin(q_5)\cos(q_4)\Big](a_5+a_6) + \sin(q_1)\cos(q_2+q_3)(a_3+a_4) + \sin(q_1)\cos(q_2)a_2 \quad (5.23)$$

$$p_z = [\sin(q_2+q_3)\cos(q_5) + \sin(q_5)\cos(q_2+q_3)\cos(q_4)](a_5+a_6) + \sin(q_2+q_3)(a_3+a_4) + \sin(q_2)a_2 + (a_0+a_1) \quad (5.24)$$

It can be seen that Eqs. (5.13, 5.14, 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, and 5.21) reflect the values of the orientation vectors of the end of the robot [*noa*] as a function of the joint coordinates $(q_1, q_2, q_3, q_4, q_5, q_6)$ and Eqs. (5.22), (5.23) y (5.24) reflect the values of the position vector of the end of the manipulator (p_x, p_y, p_z) as a function of the joint coordinates and lengths of the links $(a_0, a_1, a_2, a_3, a_4, a_5, a_6)$.

5.2.2 *Description of Data Sets*

According to the geometry and physical dimensions of the Ketzal robotic manipulator, the workspace can be represented by the set of position coordinates, orientation, and joint values. In principle, the workspace is made up of an infinite set of spatial coordinates. However, it is necessary to generate a space defined by a finite set so that it can be processed by a computer, because the volume in the data considerably influences the available processing capacities [4].

Two data sets *A* and *B* were proposed. Both sets were generated from the transformation matrices described by Eq. (5.5). According to the geometric characteristics of the Ketzal robot, the ranges of movement for each of the joints $(\Theta_1 \ldots \Theta_6)$ are described in Table 5.2.

The workspace for both proposed sets is the same; however, the distribution of the spatial coordinates is different due to the resolution used in the ranges of motion initially established in each of the joints during the generation of the two data sets. In other words, the amount of data generated is defined as a function of the spatial resolution established in the range of joint values. For example, joint Θ_1 has a range of motion from 0 to 2π; therefore, if a jump is set $\Delta\theta_1 = \pi$, only three values [0, π, 2π] are considered, on the contrary, if a jump is used from $\Delta\theta_1 = \pi/5$ a better spatial resolution is obtained considering eleven values [15].

Table 5.2 Angular ranges in the Ketzal robot joints

(rad)	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
Mínimum	0	0	2π	0	2π	0
Máximum	2π	π	$\frac{\pi}{2}$	2π	$\frac{\pi}{2}$	2π

Available hardware features and capabilities are one of the most important factors that significantly influence data processing. Below a description of the two data sets that were generated by implementing two arrays of *rows* $\times$ *columns* is presented [44].

$$\text{Dataset}A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1c} \\ a_{21} & a_{22} & \dots & a_{2c} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rc} \end{bmatrix} \tag{5.25}$$

For data set *A*, the subscript *r* represents the number of data generated by each variable with a value of 24,000 double precision data, with a spatial resolution of 10×5×5×6×4×4, that is, joint Θ_1 within its range of motion can only generate 10 values, the second joint 5 values, and so on for the other joints. The subscript *c* represents the total number of variables used with a value of 18 variables, giving a total of 432,000 data with a physical space in memory of 3.29 MB. The elements $\{a_{r1}, a_{r2}, a_{r3}\}$ correspond to the position vector $[p] = \{p_x, p_y, p_z\}$, the elements $\{a_{r4} \dots a_{r12}\}$ correspond to the orientation vector $[noa] = \{n_x, n_y, n_z, o_x, o_y, o_z, a_x, a_y, a_z\}$, and finally the elements $\{a_{r13} \dots a_{r18}\}$ correspond to the vector of joint values $[\Theta] = \{\Theta_1, \Theta_2, \Theta_3, \Theta_4, \Theta_5, \Theta_6\}$.

$$\text{Dataset}B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1c} \\ b_{21} & b_{22} & \dots & b_{2c} \\ \dots & \dots & \dots & \dots \\ b_{r1} & b_{r2} & \dots & b_{rc} \end{bmatrix} \tag{5.26}$$

For data set *B*, the subscript r has a value of 244,140,625 double precision data, with a spatial resolution of 25×25×25×25×25×25, where each of the joints can generate 10 values within its range of motion. The subscript *c* has a value of 18 variables giving a total of 4,394,531,250 data with a physical space in memory of 4.09 GB.

The elements $\{b_{r1}, b_{r2}, b_{r3}\}$ correspond to the position vector $[p] = \{p_x, p_y, p_z\}$, the elements $\{b_{r4} \dots b_{r12}\}$ correspond to the orientation vector $[n\ o\ a] = \{n_x, n_y, n_z, o_x, o_y, o_z, a_x, a_y, a_z\}$, and finally the elements $\{b_{r13} \dots b_{r18}\}$ correspond to the vector of joint values $[\Theta] = \{\Theta_1, \Theta_2, \Theta_3, \Theta_4, \Theta_5, \Theta_6\}$.

5.2.3 Data Set Collection

To collect the data sets from the kinematic analysis of the proposed manipulator, a six-dimensional matrix was generated that contains all the combinations of values defined in Θ_1, Θ_2, Θ_3, Θ_4, Θ_5, and Θ_6. For data set *B*, each of the joints was defined with 25 values; therefore, the resulting matrix has a dimension of

25x25x25x25x25x25. For data set *A*, the resulting matrix has a dimension of 10x5x5x6x4x4. The *x*, *y*, *z* coordinates are deduced through the direct kinematics equations, starting from the ranges of motion for each data set as described in Table 5.2 and taking into account the lengths of the six links of the manipulator, where $a_0 = 10.1$ cm, $a_1 = 10.1$ cm, $a_2 = 16.0$ cm, $a_3 = 9.2$ cm, $a_4 = 10.3$ cm, $a_5 = 1.345$ cm, and $a_6 = 5.37$ cm. When calculating the direct kinematics Eqs. (5.13, 5.14, 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21, 5.22, 5.23, and 5.24) as a function of the joints and lengths of the links, a six-dimensional matrix is obtained for each equation solved. Finally, to obtain the data set, the result of the direct kinematics equations is grouped together with the joint value matrix in a single matrix. The resulting matrix contains nine orientation matrices, three position matrices, and six joint value matrices, for a total of 18 six-dimensional matrices.

5.2.4 Dispersion Analysis of the Generated Data Set

Starting from the initial position and according to the range of movement defined for each of the joints, the workspace can be appreciated from different perspectives by using scatter diagrams in the position vectors as a function of the joints. Figure 5.4 shows the dispersion of the position data as a function of the three joints of the end effector, generating a half sphere as shown below.

By including the combinations of the q_1 joint with the previous one, a set of half spheres following a circular trajectory is obtained, because the joint has a range of motion from 0 to 360 degrees as seen below in Fig. 5.5.

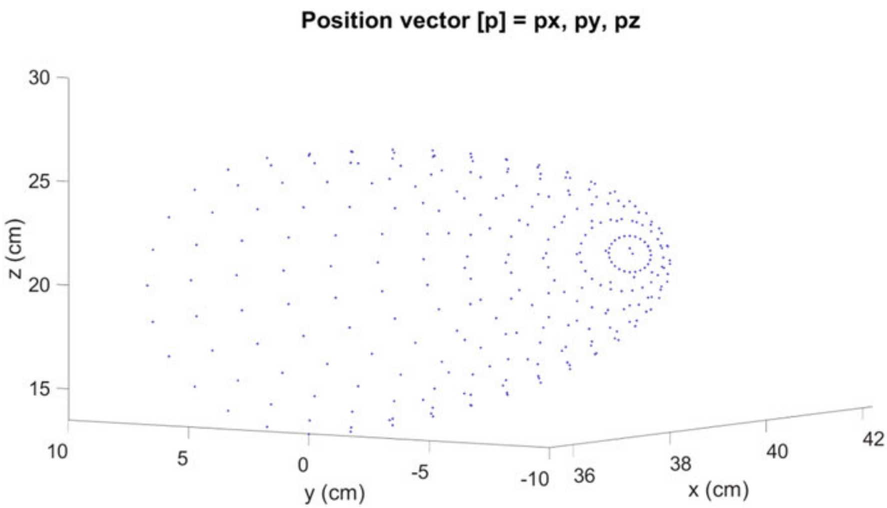


Fig. 5.4 Position vector (p_x, p_y, p_z) in function of joints q_4, q_5 **and** q_6

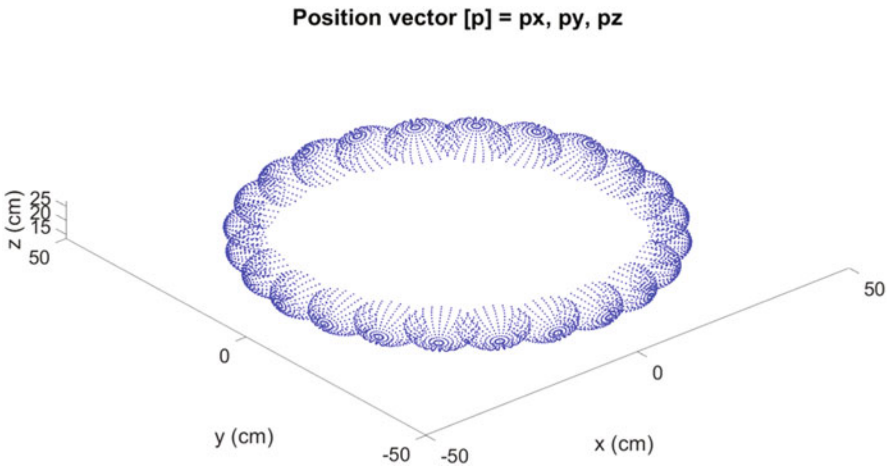


Fig. 5.5 Position vector ($\boldsymbol{p_x}, \boldsymbol{p_y}, \boldsymbol{p_z}$) in function of joints $\boldsymbol{q}_1$, $\boldsymbol{q}_4$, $\boldsymbol{q}_5$ **and** $\boldsymbol{q}_6$

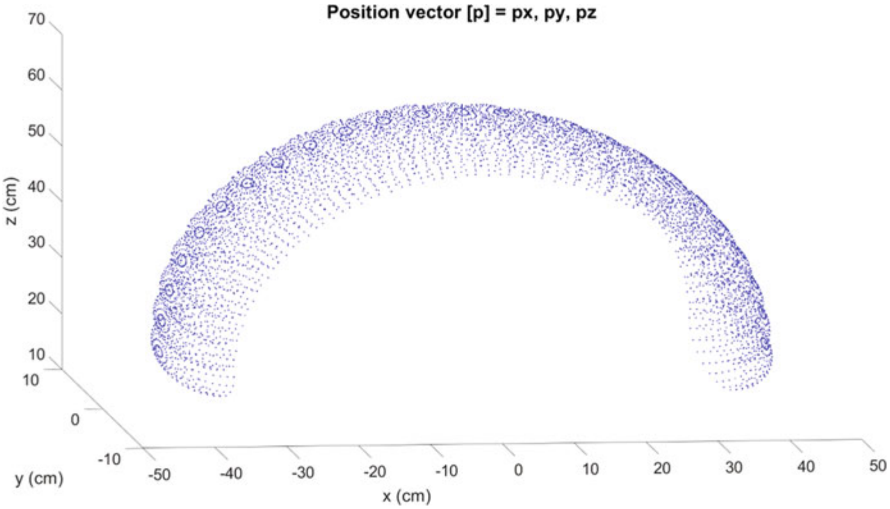


Fig. 5.6 Position vector ($\boldsymbol{p_x}, \boldsymbol{p_y}, \boldsymbol{p_z}$) in function of joints $\boldsymbol{q}_2$, $\boldsymbol{q}_4$, $\boldsymbol{q}_5$ **and** $\boldsymbol{q}_6$

On the contrary, if the combinations of the q_2 joint instead of the q_1 joint are included in the position vector functions(p_x, p_y, p_z), a set of half spheres in a vertical plane following a half-circle trajectory is obtained, because the joint q_2 has a range movement from 0 to 180 degrees as shown below in Fig. 5.6.

If the combinations of q_3, q_4, q_5 and q_6 joints are included in the position vector functions (p_x, p_y, p_z), a series of half spheres in a horizontal plane following a circular path are obtained, because the joint q_3 has a range of motion of -90 to 90 degrees as shown below in Fig. 5.7.

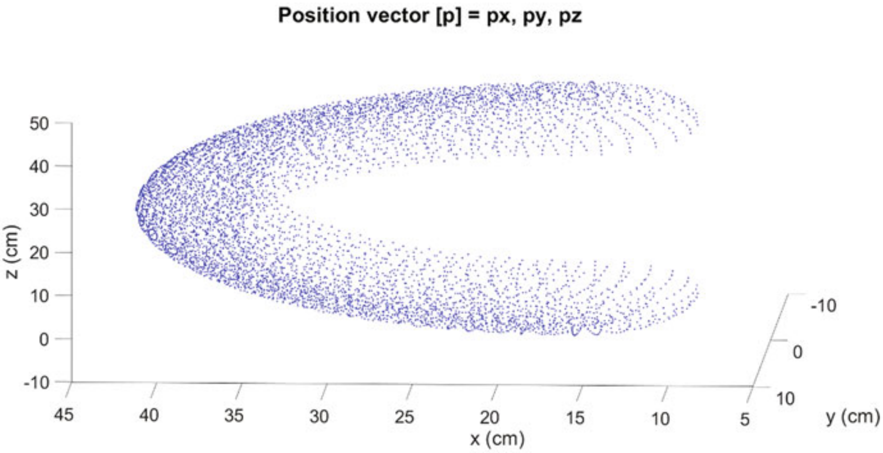


Fig. 5.7 Position vector $(\boldsymbol{p_x}, \boldsymbol{p_y}, \boldsymbol{p_z})$ in function of joints $\boldsymbol{q}_3$, $\boldsymbol{q}_4$, $\boldsymbol{q}_5$ **and** $\boldsymbol{q}_6$

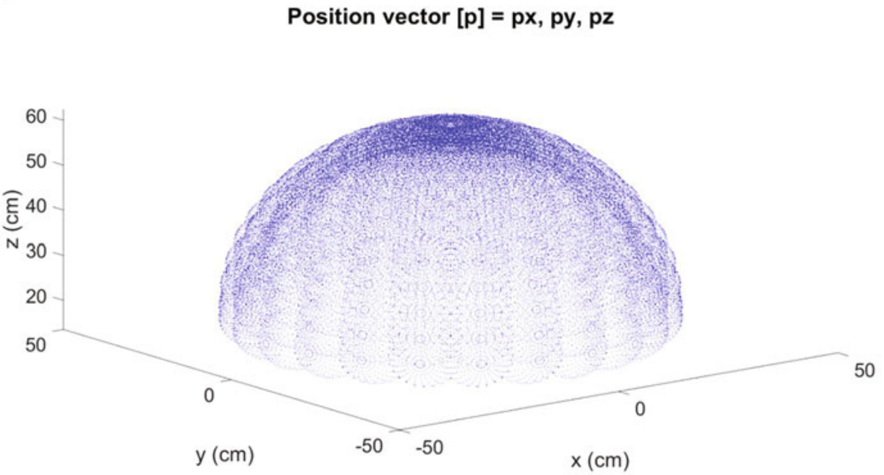


Fig. 5.8 Position vector $(\boldsymbol{p_x}, \boldsymbol{p_y}, \boldsymbol{p_z})$ in function of joints $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_4$, $\boldsymbol{q}_5$ **and** $\boldsymbol{q}_6$

In this particular case, if the q_1, q_2, q_4, q_5 and q_6 joints are combined to the position vector functions (p_x, p_y, p_z), a vertical and horizontal path of the half sphere generated by the q_4, q_5 and q_6 joints is obtained through a half-circle path as shown below in Fig. 5.8.

Finally, by combining all the joints of the manipulator, a data set of size 244,140,625 data with 18 variables is obtained and generated through the functions of the position vector (p_x, p_y, p_z) as shown below in Fig. 5.9.

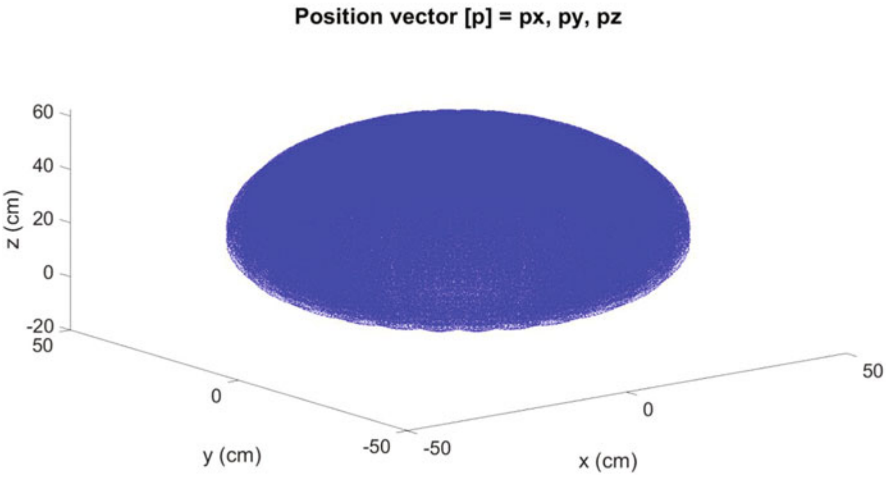


Fig. 5.9 Position vector (p_x, p_y, p_z) in function of joints q_1, q_2, q_3, q_4, q_5 **and** q_6

5.2.5 Reduction Algorithm Based on Linear Systematic Sampling

It is evident that the amount of data generated in set B exceeds the processing capacities of a conventional computer. To solve the problem, an algorithm based on the LSS method was designed. This method has two main advantages. On one hand, the selection of the first sample is chosen by a sampling period that guarantees a random instance among a set of samples in a given interval. On the other hand, the samples chosen consecutively are distributed in a good way among the population, that is, there is less risk that some or a large part of the population is not represented, maintaining a constant and uniform distribution among the data [4].

Although the LSS method is one of the most common methods, it has two drawbacks. On one hand, The sampling variance cannot be taken impartially on the basis of the single sample taken and in the other hand, when the population size N cannot be divided evenly by the desired sample size n, systematic sampling cannot be performed, that is, when N is not an integer multiple of the desired sample size n and consequently $N \neq nk$. In this case, the sampling could be inefficient and if at some point the characteristics of the population were periodic and they coincide with the sampling interval, the representativeness of the desired sample could be biased [10, 11].

The quality of the data is influenced by the spatial resolution of the data in the workspace. In this sense, the higher the spatial resolution, the more data will be distributed in the workspace and the better the data quality. However, to carry out the processing on conventional processors, it will be necessary to implement data reduction algorithms that allow obtaining a balanced representation and thus obtaining good results in the knowledge extraction algorithms [4, 6].

In this case, considering data set B with a population of $N = 244{,}140{,}625$, where the size of the desired sample is $n = 24{,}000$, it is observed that N is not an integer multiple of n, because $k = N / n = 10{,}172.5260$; therefore, k is not an integer that satisfies the criteria of the LSS method. However, due to the distribution characteristics of the data, it is possible to carry out systematic sampling while maintaining a homogeneous representativeness among the data. In this case, the sampling interval k will be converted to an integer, because the characteristics of the population are not periodic; therefore, the accumulated bias at the end of the sampling can be eliminated without altering or modifying the homogeneous representativeness of the population. Figure 5.10 describes the RDF algorithm that was used only with data set B due to its size to fulfill this purpose.

According to the Fig. 5.10, the steps for the implementation of the filter are described in a general way:

1. The values for the original population size N and the desired sample size n are initialized.
2. The constant of the sampling interval $k = N / n$ is calculated.
3. The random number between 1 and k is generated.
4. The random value is rounded to an integer *r1*.
5. Element *r1* from the original data set is selected as the first filtered data.

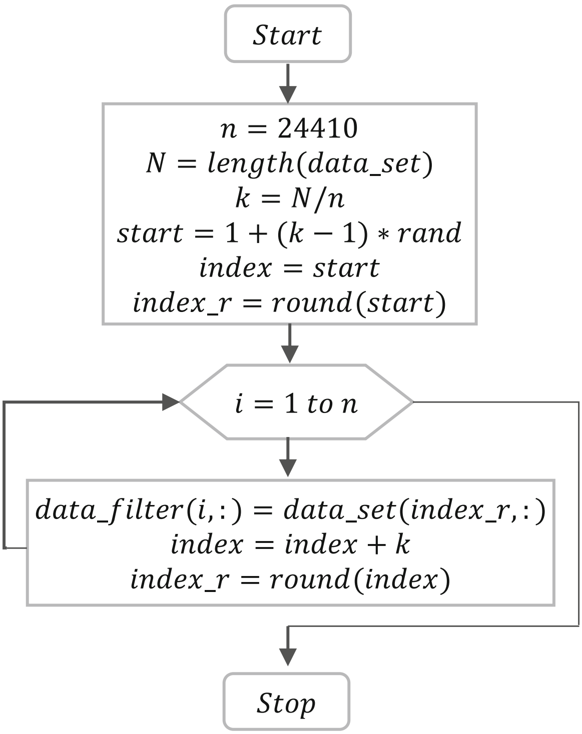


Fig. 5.10 RDF algorithm

6. The cumulative sum is made between the random number and the constant k. In this step, the cumulative sum is a positive noninteger value.
7. The cumulative value of the sum is rounded to an integer $r2$.
8. Element $r2$ is selected from the original data set as the second filtered data and steps 6, 7, and 8 are repeated until the desired n samples are reached.

Where the variable n represents the desired sample, N represents the original population, k is the sampling interval, a random number between 1 and k is stored in the variable *start*, and the variable *index* represents the index of the instance with decimal number while *index_r* is the integer value of the index of the instance rounded up.

Within the iterative loop type for, the selection of instances is performed starting from the *k-th* element until reaching the desired sample n. The variable index accumulates the values of the indices with decimals and the value of the sampling interval k.

For the case of data set B, there is a population $N = 244{,}140{,}625$, where the desired sample is $n = 24{,}414$ and the sampling interval $k = 10{,}000{,}025$. In this case, when k is not an integer, the RDF algorithm will always perform rounding. Therefore, when the value of the modulus of the sampling interval k divided by the desired sample n is zero, a sampling without bias among the population is ensured, that is, when K is an integer, the LSS criterion is applicable. However, when the value of the modulus of the sampling interval divided between the desired samples is close to zero, the bias generated is minimal without significantly influencing the training results.

5.2.6 Data Set Normalization

The normalization of data sets A and B was performed in the range of -1 to 1, because the hyperbolic sigmoid tangent transfer function was used for the BPANN training. A mean of zero was applied, in an interval of $-1 \leq \text{data} \leq 1$ by means of Eq. (5.27).

$$\text{datanorm} = \left(\frac{\text{data-}\min}{\text{range}} * (\text{high} - \text{low})\right) + \text{low} \tag{5.27}$$

where *data_norm* is the normalized data, *data* is the data to normalize, *min* is the minimum of the generated data set, *range* is the range or difference between the maximum value and the minimum value of the generated data set, $high = 1$ represents the upper limit of the desired interval, and $low = -1$ represents the lower limit of the desired range [45].

5.2.7 Training and Test Data Sets

There is no procedure that determines the amount of training and test data that should be used to train ANNs. However, much of the research with ANN uses the 80:20 and 90:10 ratio. In this study, a random selection of the data was carried out in order to randomly choose the training and testing subsets with densities of 80:20 and 90:10. This procedure was applied to both data sets, *A* and *B*.

To analyze the performance in the proposed network architectures, a training was carried out for each proposed data density configuration. With the aim of analyzing the generalizability in each of the proposed network architectures, during the testing stage, an error of less than 5% was considered for the prediction of the data through a statistical analysis of correlation and Chi-square.

5.2.8 Training and Test Back Propagation Neural Network

To determine the optimal parameters of the BPNN network, the RDANN methodology based on the robust parameter design method proposed by Genichi Taguchi was used. The engineering method, applied to the design of products or processes, is focused on reducing the sensitivity to noise and has proven to be an efficient and powerful method in product design [46].

The robust design technique is known as factorial design of experiments where most of the possible combinations can be identified without the need to include a considerable number of experiments and its application allows determining the functionality or performance of a product or process to be controlled [47].

The RDANN methodology applied to the ANN design allows finding the selection of the factors involved that allow minimizing the variability of the response to different inputs to the system through the appropriate choice of the levels in the controlled design variables [19].

The design variables considered were the number of neurons in the first and second layers, the momentum, and the learning rate. For the noise variables, the random weights, the size of the training set versus the size of the test set, and the random selection of the sets were considered. During the experimentation stage, an orthogonal array configuration was used with an $L_9(3^4)$ y $L_4(3^2)$ with the aim of training and testing 36 different ANN architectures [48].

During the confirmation stage, the signal-to-noise ratio was analyzed through an analysis of variance (ANOVA) to determine adequate levels in the variables involved and to identify the possible optimal values of the best network topology, also involving the value of the mean obtained from the Mean Square Error (MSE). The best architecture was 12: 12: 6, momentum = 0.01, and learning rate = 0.1.

Figure 5.11 shows the generalization tests applied to the BPNN that was trained with 19,200 data from set A with spatial resolution of $10 \times 5 \times 5 \times 6 \times 4 \times 4$ without applying the filter. In each individual graph, the six joint coordinates of

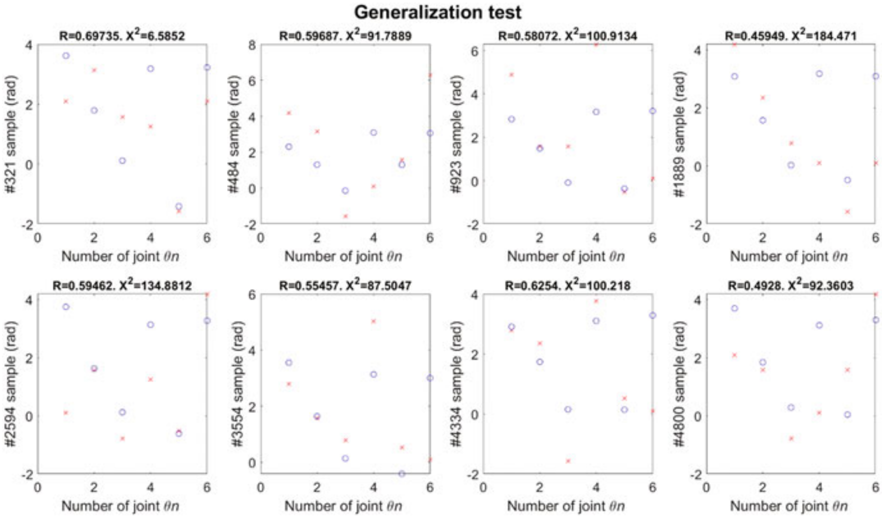


Fig. 5.11 Without filtering: Predicted joints vs calculated

the manipulator that were previously calculated ("x") versus the joint coordinates predicted by the BPNN ("o") are observed.

Figure 5.12 shows the tests carried out on the BPNN, with the data from set *B* that were previously treated by the RDF, with a spatial resolution of 25×25×25×25×25×25. Joint coordinates previously calculated and belonging to the test data set are displayed. In each box, six joint coordinates calculated ("x") versus joint coordinates predicted by the BPNN ("o") are observed.

5.2.9 Training and Test Generalized Regression Neural Network

In the GRNN, to determine the optimal spread value of the network, also known as the "kernel spread constant," a spread value smaller than the distance between the input vectors is used to make a close fit between the data. If a higher spread value is used, it could cause an over adjustment, on the contrary, if the value is too small, it could cause an under adjustment. Therefore, this constant propagation value is considered as a regularization parameter that should be optimally selected [49].

To determine the best constant propagation value in the RGNN, 2000 trainings were performed for each configuration in the data density (80:20 and 90:10) for each proposed set (A and B) in an automated manner. Next, Table 5.3 shows the spread values obtained during the RGNN training.

Figure 5.13 shows the generalization tests applied to the GRNN that was trained with 19,200 data from set A with spatial resolution of 10×5×5×6×4×4, without

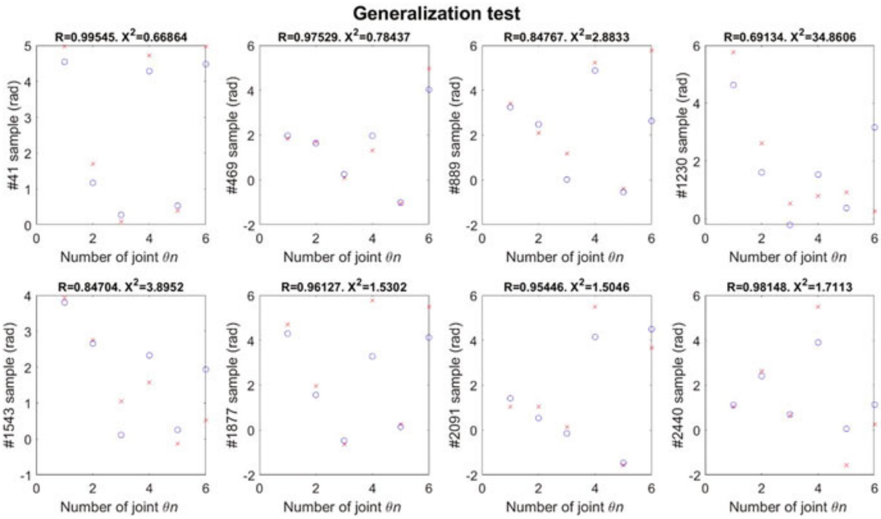


Fig. 5.12 With filtering: Predicted joints vs calculated

Table 5.3 Obtaining the spread value with 2000 iterations

Data set	Train/test	Global time [hr]	Optimal spread
A	80:20	2.38	0.2881
A	90:10	1.22	0.2721
B	80:20	2.26	0.2111
B	90:10	0.89	0.2131

applying the filter, with a much lower spatial resolution compared to set *B*. In each individual graph, the six joint coordinates of the manipulator that were previously calculated ("x") versus the joint coordinates predicted by the GRNN ("o") are observed.

Figure 5.14 shows the tests performed on the GRNN that was trained with the data from set B that were previously treated by the RDF, with a spatial resolution of 25x25x25x25x25x25. Eight joint coordinates previously calculated and belonging to the test data set are displayed. In each box, six joint coordinates calculated ("x") are observed versus joint coordinates predicted by the GRNN ("o") are observed.

5.3 Results

5.3.1 Reduction Data Filter Analysis

To analyze the distribution of the data, two dispersion matrices were plotted to compare the results obtained before and after applying the filter to the data. Based on the range of motion previously established for each of the joints in Fig. 5.15 (a), a

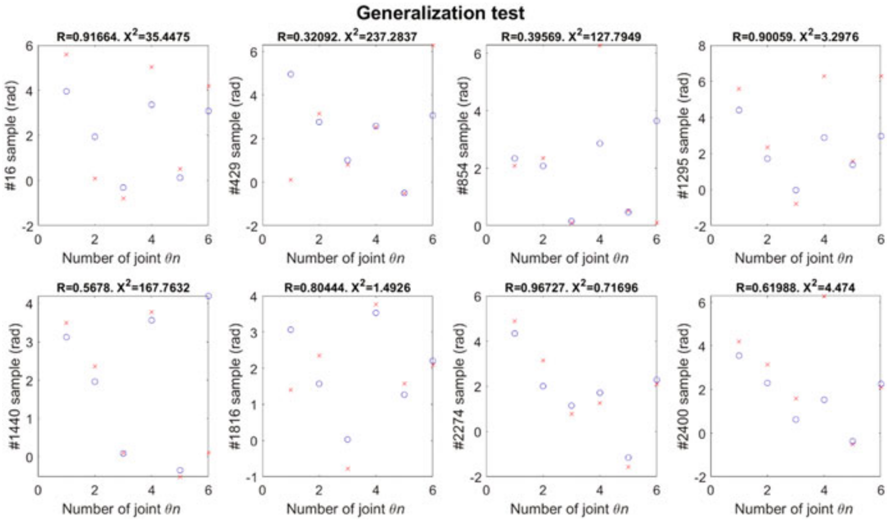


Fig. 5.13 Without filtering: Predicted joints vs calculated

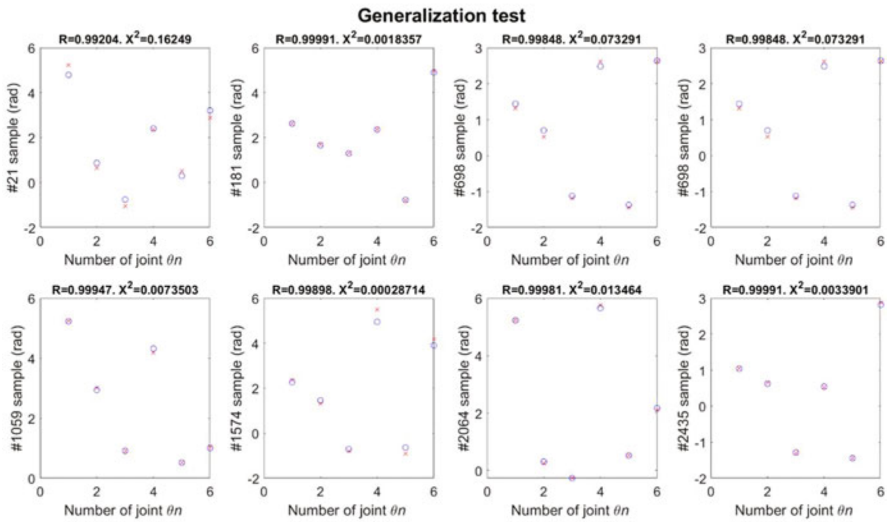
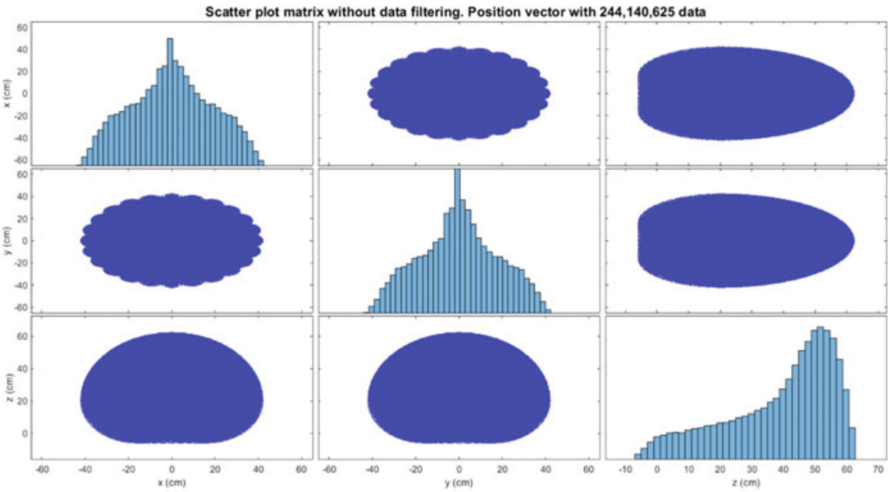
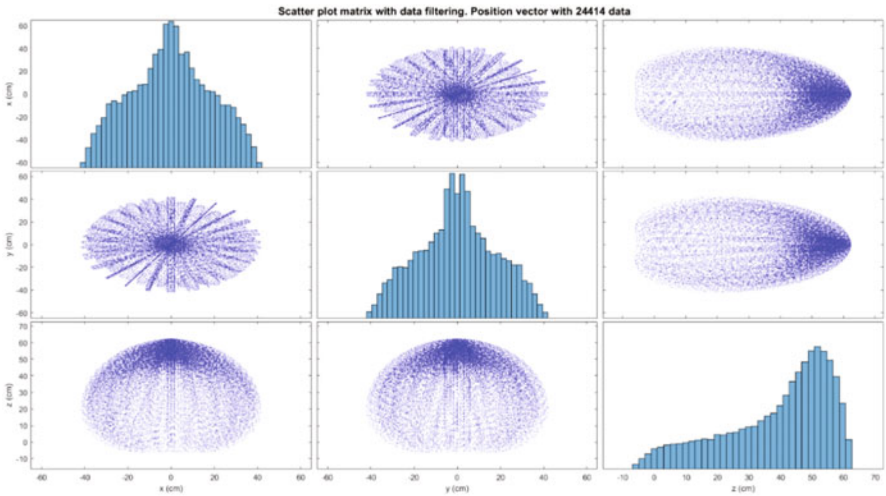


Fig. 5.14 With filtering: Predicted joints vs calculated

representation of data set B is shown, which maintains a population of 244,140,625 data for each variable. On the main diagonal, the distribution of three variables corresponding to the position vector $[p] = \{p_x, p_y, p_z\}$ and their data scatter diagrams in different perspectives can be observed.



(a)

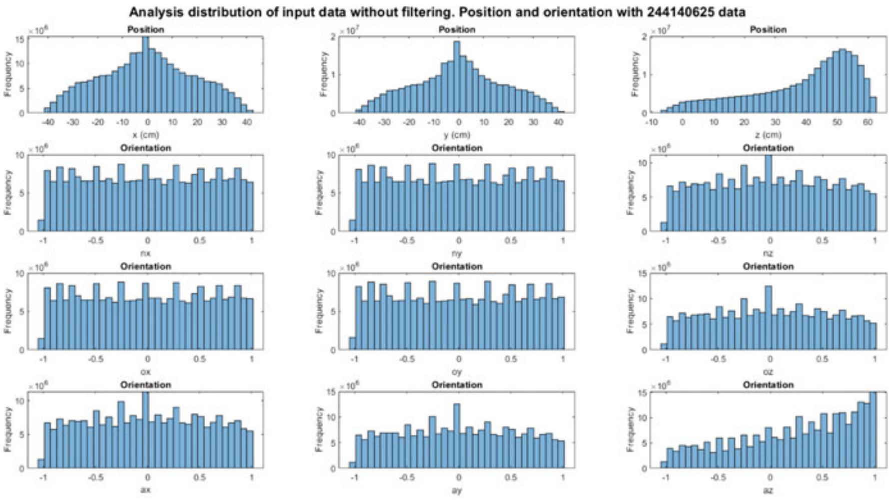


(b)

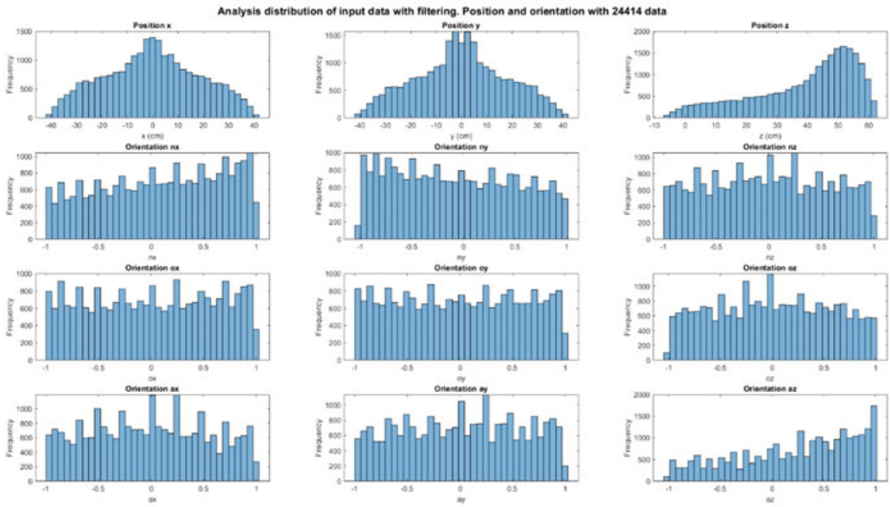
Fig. 5.15 Dispersion matrix of the position data set *B*: (**a**) Before filtering (**b**) After filtering

In Fig. 5.15 (b), the data after applying the RDF are shown, obtaining a reduction of $n = 24{,}410$ data for each variable. It can be seen that the data maintain a constant and uniform distribution with respect to the data set shown in Fig. 5.15 (a).

Figure 5.16 shows the distribution of the position and orientation data vectors before and after applying the RDF to the data. In Fig. 5.16 (a), we can see the distribution of the vectors of position $[p] = \{p_x, p_y, p_z\}$ and orientation $[n\ o\ a] = \{n_x,$



(a)



(b)

Fig. 5.16 Distribution of position and orientation data set *B*: (**a**) Before filtering (**b**) After filtering

n_y, n_z, o_x, o_y, o_z, a_x, a_y, a_z} belonging to data set B and which originally maintains a population of 244,140,625 data per variable.

In Fig. 5.16 (b), the data distribution after applying the RDF is shown, obtaining a reduction of $n = 24{,}410$ data per variable. It can be seen that most of the data maintains the distribution in a constant and uniform way with respect to the data set shown in Fig. 5.16 (a).

5.3.2 *Performance Evidence with Filtering: Comparison GRNN with BPNN*

Next, Table 5.4 shows the results obtained by the two network architectures in the testing stage, where it can be seen at the end of the table that RGNN obtained the best percentage of predictions with an error of less than 5%. Likewise, it can be observed that the training time required for this network is much lower compared to the BPNN.

The use of the RDF filter allowed solving the problem of data volume and also to carry out the training in a conventional processor, maintaining the homogeneous distribution of the original data. During the tests, an error of less than 5% was considered in the prediction of the data for both architectures by means of a statistical analysis of correlation and Chi square.

For the BPNN training, 10,000 iterations were performed and the *trainrp* training algorithm was used with a goal mse = 1E-4. The structural parameters of the network such as momentum and learning rate were obtained by the RDANN methodology with values of 0.01 and 0.1, respectively. The best results in this architecture were obtained through the training performed by data set B processed by the RDF Filter, with a data density of 90% for training data and 10% for tests.

5.4 Conclusions and Discussions

There is a relationship between the improvement in the generalizability of both proposed networks and the fact that the data set was previously processed by the RDF, because the reduction of the data allows maintaining a representative distribution with respect to the spatial resolution generated initially, maintaining the characteristics of the original population, and obtaining a considerable reduction of 99.99% to be able to carry out the processing in conventional computers and obtain good results.

For the application of the RDF algorithm, as long as the modulus of the sampling interval k divided by the desired sample n is a value close to zero, it will be possible to obtain a better representativeness of the population N. Therefore, the answer of question How close to zero should the sampling interval k divided by the desired sample n be? its determination will depend on the characteristics of the problem to consider it as a sampling that allows obtaining good results, since the expected results are not always obtained when considering an integer k sampling interval. In this study, the value was 0.02560.

The results obtained by the two network architectures showed an increase in the generalizability due to the fact that they were trained with the data set previously treated by the filter. However, when training both architectures using the data set that was not treated, the generalizability in both networks was lower.

Table 5.4 Comparative table of results obtained

				RGNN		BPNN	
Data set	Size before Filtering	Size after filtering	Train/Test	Training time [s]	%Rightguess $X^2 < 5\%$	Training time [s]	%Rightguess $X^2 < 5\%$
A	24,000 × 18	24,000 × 18	80:20	4.27	48%	264.41	44%
A	24,000 × 18	24,000 × 18	90:10	1.60	50%	316.02	46%
B	244,140,625 × 18	24,000 × 18	80:20	7.41	79%	175.44	71%
B	244,140,625 × 18	24,000 × 18	90:10	3.51	80%	184.84	72%
B	244,140,625 × 18	24,414 × 18	80:20	2.69	83%	189.04	73%
B	*244,140,625 × 18*	*24,414 × 18*	*90:10*	*1.58*	*83%*	*158.57*	*74%*

The increase in the percentage of successes in relation to the spatial resolution in the training data sets is evident. However, as spatial resolution increases, data processing requires higher performance in the available hardware resources, so the benefit that the use of data preprocessing techniques and/or tools that allow obtaining good results without decreasing performance and generalizability in knowledge extraction algorithms is evident.

The time required to train a BPNN is usually ten times greater than the training time required by a GRNN, in addition to the fact that it is not necessary to select a list of structural parameters compared to the BPNN, so its implementation is faster and more efficient.

The best training in both network architectures was obtained using a data density of 80% for training and 20% for testing. At best, during the testing phase, the GRNN succeeded in keeping 83% of its predictions with a margin of error of less than 5%. Therefore, this study reaffirms that the quality of the training data sets has a significant influence on the results obtained by the knowledge extraction algorithms.

Acknowledgments This work was supported by IPN-COTEBAL to study doctoral degree under permissions 14/2018, 33/2019, and 23/2020. This works was partially supported by CONACYT - BecasNacionales de Posgrado con la Industria under contract 431101/640582, by the collaborate students Agustín Ortíz and Brayan M. Carrera and by OMADS S.A. of C.V., an enterprise dedicated to the innovation and technological development.

References

1. C. Marco, P. Anit, How organisations leverage big data: A maturity model. Ind. Manag. Data Syst. **116**(8), 1468–1492 (2016). https://doi.org/10.1108/IMDS-12-2015-0495
2. K.-C. Li et al., *Big Data: Algorithms, Analytics, and Applications* (Chapman and Hall/CRC, 2015)
3. H. Liu, H. Motoda. On issues of instance selection. Data Min. Knowl. Disc. **6**(2) Art. no. 2 (2002). https://doi.org/10.1023/A:1014056429969
4. H. Liu, H. Motoda, Instance Selection and Construction for Data Mining. Springer US (2001)
5. S. García, J. Luengo, F. Herrera, *Data Preprocessing in Data Mining*, vol 72 (Springer International Publishing, Cham, 2015)
6. B. Gu, F. Hu, H. Liu. Sampling and its application in data mining. Technical Report TRA6/00, Department of Computer Science, National University of Singapur (2000)
7. H. Liu, H. Motoda, On issues of instance selection. Data Min. Knowl. Disc. **6**(2), 115–130 (2002). https://doi.org/10.1023/A:1014056429969
8. H. Brighton, C. Mellish, Advances in instance selection for instance-based learning algorithms. Data Min. Knowl. Disc. **6**(2), 153–172 (2002). https://doi.org/10.1023/A:1014043630878
9. W. G. Madow, L. H. Madow. On the theory of systematic sampling, I. Ann. Math. Stat. **15(** 1). Art. no. 1 (1944)
10. S.A. Mostafa, I.A. Ahmad, Remainder linear systematic sampling with multiple random starts. J. Statist. Theory Pract. **10**(4), 824–851 (2016). https://doi.org/10.1080/15598608.2016.1231094
11. L.H. Madow, Systematic sampling and its relation to other sampling designs. J. Am. Stat. Assoc. **41**(234), 204–217 (1946). https://doi.org/10.1080/01621459.1946.10501864
12. A. Serrano García. Inteligencia artificial (2016)

13. X. He, S. Xu, SpringerLink (Online service), *Process Neural Networks: Theory and Applications* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010)
14. A.K. Jain, J. Mao, K.M. Mohiuddin, Artificial neural networks: A tutorial. Computer **29**(3), 31–44 (1996). https://doi.org/10.1109/2.485891
15. L. Aggarwal, K. Aggarwal, R.J. Urbanic, Use of artificial neural networks for the development of an inverse kinematic solution and visual identification of singularity zone(s). Procedia CIRP **17**, 812–817 (2014). https://doi.org/10.1016/j.procir.2014.01.107
16. J. Zupan, *Introduction to Artificial Neural Network (ANN) Methods: What They Are and How to Use Them*, vol 41 (1994)
17. Y. Zhang, D. Guo, Z. Li, Common nature of learning between back-propagation and hopfield-type neural networks for generalized matrix inversion with simplified models. IEEE Trans Neural Netw. Learn. Syst. **24**(4), 579–592 (2013). https://doi.org/10.1109/TNNLS.2013.2238555
18. T. Ozaki, T. Suzuki, T. Furuhashi, S. Okuma, Y. Uchikawa, Trajectory control of robotic manipulators using neural networks. IEEE Trans. Ind. Electron. **38**(3), 195–202 (1991). https://doi.org/10.1109/41.87587
19. J.M. Ortiz, R. del Martínez, J.M.C. Viramontes, H.R. Vega, *Robust Design of Artificial Neural Networks Methodology in Neutron Spectrometry* (Artificial Neural Networks - Architectures and Applications, 2013). https://doi.org/10.5772/51274
20. Specht, Probabilistic neural networks for classification, mapping, or associative memory, in *IEEE 1988 International Conference on Neural Networks*, 1988, pp. 525–532 vol.1, doi: https://doi.org/10.1109/ICNN.1988.23887
21. D.F. Specht, Probabilistic neural networks. Neural Netw. **3**(1), 109–118 (1990). https://doi.org/10.1016/0893-6080(90)90049-q
22. D.F. Specht, P.D. Shapiro, Generalization accuracy of probabilistic neural networks compared with backpropagation networks. IJCNN-91-Seattle Int. J. Conf. Neural. Netw. **i**, 887–892 (1991). https://doi.org/10.1109/IJCNN.1991.155296
23. P. Jha, B. B. Biswal. A neural network approach for inverse kinematic of a SCARA manipulator. *IAES International Journal of Robotics and Automation*, **3**(1), Art. no. 1 (2014). https://doi.org/10.11591/ijra.v3i1.3201
24. Lee, Robot arm kinematics, dynamics, and control. Computer **15**(12), 62–80 (1982). https://doi.org/10.1109/MC.1982.1653917
25. B. Karlik, S. Aydin, An improved approach to the solution of inverse kinematics problems for robot manipulators, Eng. Appl. Artif. Intell. **13**(2), Art. no. 2, (2000). https://doi.org/10.1016/S0952-1976(99)00050-0
26. L. Jin, S. Li, J. Yu, J. He, Robot manipulator control using neural networks: A survey. Neurocomputing **285**, 23–34 (2018). https://doi.org/10.1016/j.neucom.2018.01.002
27. S. Li, Y. Zhang, L. Jin, Kinematic control of redundant manipulators using neural networks. IEEE Transact Neural Netw Learn Syst **28**(10), 2243–2254 (2017). https://doi.org/10.1109/TNNLS.2016.2574363
28. M. Tarokh, M. Kim, Inverse kinematics of 7-DOF robots and limbs by decomposition and approximation. IEEE Trans. Robot. **23**(3), 595–600 (2007). https://doi.org/10.1109/TRO.2007.898983
29. R. Köker, T. Çakar, Y. Sari, A neural-network committee machine approach to the inverse kinematics problem solution of robotic manipulators. Eng. Comput. **30**(4), 641–649 (2014). https://doi.org/10.1007/s00366-013-0313-2
30. B.K. Bose, Neural network applications in power electronics and motor drives—An introduction and perspective. IEEE Trans. Ind. Electron. **54**(1), 14–33 (2007). https://doi.org/10.1109/TIE.2006.888683
31. A. Hasan, A.T. Hasan, H.M.A.A. Al-Assadi, Performance prediction network for serial manipulators inverse kinematics solution passing through singular configurations. Int. J. Adv. Robot. Syst. **7**(4), 11–24 (2011)
32. R. Gao. Inverse kinematics solution of Robotics based on neural network algorithms. J. Ambient Intell. Humanized Comput. (2020) https://doi.org/10.1007/s12652-020-01815-4

33. L. Jin, S. Li, J. Yu, J. He, Robot manipulator control using neural networks: A survey. Neurocomputing **285**, 23–34 (2018). https://doi.org/10.1016/j.neucom.2018.01.002
34. A.-M. Zou, Z.-G. Hou, S.-Y. Fu, and M. Tan. Neural Networks for Mobile Robot Navigation: A Survey. In *Advances in Neural Networks - ISNN 2006* (2006), pp. 1218–1226
35. X. Wu, Z. Xie, Forward kinematics analysis of a novel 3-DOF parallel manipulator. Scientia Iranica. Transact. B Mechan. Engin. **26**(1), 346–357 (2019). https://doi.org/10.24200/sci.2018.20740
36. V. Khoshdel, A. Akbarzadeh, Eds. An optimized artificial neural network for human-force estimation: consequences for rehabilitation robotics. Industr. Robot. Intern. J. **45**(3), Art. no. 3 (2018). https://doi.org/10.1108/IR-10-2017-0190
37. R. Fernando, Robótica – control de robots manipuladores. Alfaomega Grupo Editor (2011)
38. R. Köker, Reliability-based approach to the inverse kinematics solution of robots using Elman's networks. Eng. Appl. Artif. Intell. **18**(6), 685–693 (2005). https://doi.org/10.1016/j.engappai.2005.01.004
39. A. Larrañaga. 3D Printable Robotic Arm. GitHub (2018). https://github.com/AngelLM . Accessed 18 Sep 2019
40. R. Köker, C. Öz, T. Çakar, H. Ekiz, A study of neural network based inverse kinematics solution for a three-joint robot. Robot. Auton. Syst. **49**(3), 227–234 (2004). https://doi.org/10.1016/j.robot.2004.09.010
41. S. Tejomurtula, S. Kak, Inverse kinematics in robotics using neural networks. Inf. Sci. **116**(2), 147–164 (1999). https://doi.org/10.1016/S0020-0255(98)10098-1
42. J. Denavit, R.S. Hartenberg, A kinematic notation for lower-pair mechanisms based on matrices. Trans. ASME J. Appl. Mech. **22**, 215–221 (1955)
43. A.R.J. Almusawi, L.C. Dülger, S. Kapucu, A new artificial neural network approach in solving inverse kinematics of robotic arm (Denso VP6242). Comput Intell Neurosci CIN **2016** (2016). https://doi.org/10.1155/2016/5720163
44. S. García, J. Luengo, F. Herrera, *Data Preprocessing in Data Mining*, vol 72 (Springer International Publishing, Cham, 2015)
45. T. Jayalakshmi, A. Santhakumaran, Statistical normalization and Back propagation for classification. Intern. J. Comput. Theory Eng. **3**(1), 89–93 (2011)
46. J. Limon-Romero, D. Tlapa, Y. Baez-Lopez, A. Maldonado-Macias, L. Rivera-Cadavid, Application of the Taguchi method to improve a medical device cutting process. Int. J. Adv. Manuf. Technol. **87**(9–12), 3569–3577 (2016). https://doi.org/10.1007/s00170-016-8623-3
47. M. Ibrahim, N. Zulikha, Z. Abidin, N.R. Roshidi, N.A. Rejab, M.F. Johari, Design of an Artificial Neural Network Pattern Recognition Scheme Using Full Factorial Experiment. Appl. Mech. Mater. **465–466**, 1149–1154 (2013). https://doi.org/10.4028/www.scientific.net/AMM.465-466.1149
48. T. Y. Lin, C. H. Tseng. Optimum design for artificial neural networks: an example in a bicycle derailleur system. Eng. Appl. Artif. Intell. **13**(1), Art. no. 1 (2000). https://doi.org/10.1016/S0952-1976(99)00045-7
49. D.-S. Huang, Radial basis probabilistic neural networks: model and application. Int. J. Pattern Recognit. Artif. Intell. **13**(7), 1083–1101 (1999). https://doi.org/10.1142/S0218001499000604

Chapter 6
Smart Farming Prediction System Embedded with the Internet of Things

R. Mallikka, S. S. Manikandasaran, and K. S. Karthick

6.1 Introduction

Recently, a lack of water for agriculture has become a rising concern, especially for Asian countries such as India or Mediterranean countries. The Mediterranean countries are the most helpless against the dry season among the countries in Europe [1]. A good irrigation system is a basic requirement for the fortitude of farmers because it gives water—i.e., the lifeblood of crops—to the growing plants. Irrigation systems include various methods of getting water from various sources and conveying it to different mediums and are essential in all types of agriculture because of the unpredictable nature of the weather. It is contended that understanding the feasible and cost-proficient conventional water system techniques is important for local networks in India. The three regular conventional strategies that exist in India are complex channels utilization of stones and tree limbs, small-scale water bodies such as water system tanks to store water, and wells to gather groundwater. These techniques are commonly intended for small scale/local use for a town and large-scale or territorial applications. The small-scale customary water system techniques are not only arranged and developed by the local individuals but are also overseen by them locally, while the large-scale conventional water system

R. Mallikka (✉)
Department of Computer Science, Engineering and Applications, Bharathidasan University, Trichy, Tamilnadu, India

S. S. Manikandasaran
PG & Research, Department of Computer Science, Adaikalamatha College, Vallam, Thanjavur, Tamilnadu, India

K. S. Karthick
Bharath Niketan Engineering College, Theni, Tamilnadu, India

S. Paul et al. (eds.), *Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-77558-2_6

strategies include the surface or flood water system techniques such as bowl, fringe, and wrinkles.

The plan and structure of every conventional water system technique are chosen by the territory and precipitation history of the district [2]. In regions of good groundwater springs, burrowed wells with creative techniques are utilized to lift the water to the surface, and in the low precipitation regions, individuals utilize strategies that capture every single drop of water where it falls. These traditional water system techniques have made life possible, even in the desert areas.

The Internet of Things (IoT) speaks to a general idea for the capacity of system gadgets to detect and gather information from our general surroundings, and share that information over the Internet where it tends to be prepared and used for different fascinating purposes. The Internet of Things is rapidly turning into a reality. We can see the evidence of it around us. The Internet of Things is a system of physical items that interface with one another through the web. Things or objects can move data remotely without requiring human collaboration.

A farmer from Bahirwadi [3], (Vishwanath) a town in the dry spell inclined Beed area of Maharashtra, has earned Rs 7 lakh (700,000 rupees; ~9592 USD) from farming on only one acre of land. This farmer chose to try multi-cropping, and he likewise decided that he could expand his harvest by building a wire fence and planting creepers and climbers on them. He additionally introduced a pipeline with his first-year earnings to guarantee sprinklers watered his plants. He has also used cultivating strategies such as raised-bed gardening and mulching throughout the year, which has been useful.

In this work, we combined a smart irrigation prediction system with the IoT to improve farming. The challenges of the existing methods are a lack of groundwater and rain, which leads to an insufficient food supply for society. To overcome the problems of the exiting methods used in irrigation systems, our proposed method used various sensors for temperature, humidity, and soil moisture. Farmers are able to predict market value based on suitable crops, soil type, appropriate weather, water requirements, maintenance cost per acre, and analysis of farming risk factors.

The researchers in [4] have recommended farming systems dependent on the embedded frameworks, IoT, and remote sensor networks for agri-farm fields and domesticated animal ranches. This chapter incorporates the portrayal of frameworks with the electronic hardware of the frameworks, utilized organization conventions, and smart remote monitoring frameworks for PCs and smartphones, and so forth. Social IoT or SIoT [5] can hugely contribute to help gauge climate, pest control, humidity, precipitation, soil fertility, determine leaf moisture levels, temperature variation, airflow, and soil moisture. SIoT can likewise control plant ecological variables through temperature, moisture levels, carbon dioxide focus, and brightening as indicated by the state of crop development in real-time, and it tends to be utilized in vertical farming.

Farming 4.0 is a term for following the enormous patterns confronting the industry, helping to bring exactness to the agribusiness, using the IoT and big data to drive business efficiency despite the rising population and environmental change. The researchers in [6] have described consolidating machine learning calculations to the embedded hardware stage for information examination to help farmers achieve precision farming. The model utilizes Raspberry Pi. Machine learning calculations have been embedded for data acquisition from different sensors such as pH sensors, fire sensors, and pressure sensors. This examination inspects SVM usage in embedded processor models and offers improved engineering productivity. The current agricultural robots are being leased, sold or enlarging social work [7] on the terrains, where products of the soil are developed. With the utilization of robots in farming, the yield of production can be increased, while the expenses related to production can be decreased over time.

The rest of the chapter is organized as follows: Section 6.2 covers the background study about the smart farming system, Sect. 6.3 explains traditional methods for smart irrigation systems, Sect. 6.4 discusses the proposed irrigation prediction system, Sect. 6.5 explain the results, and Sect. 6.6 concludes the chapter and mentions the future direction of the work.

6.2 Background Study

The taxonomy of the traditional farming system is shown in Fig. 6.1, and it includes categories such as farming, agriculture, and crops with various methods of using the tradition farming system.

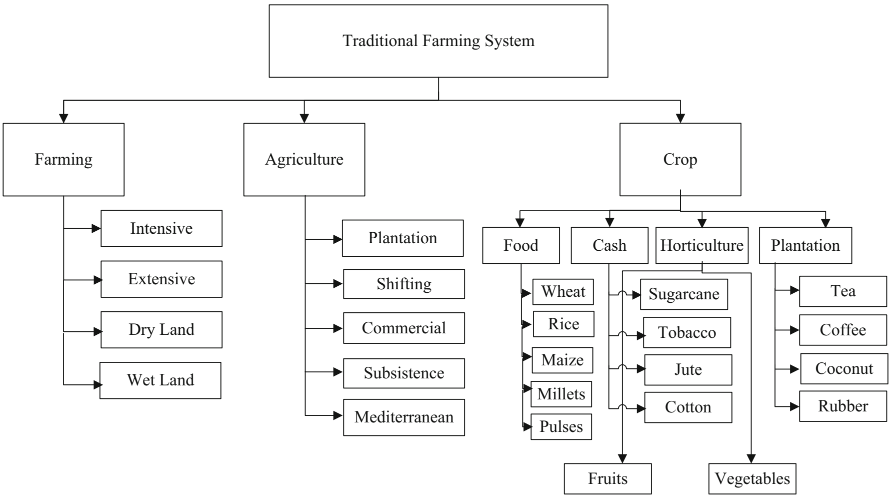


Fig. 6.1 Taxonomy of the traditional farming system

6.2.1 Traditional Farming System

The authors in [8] provided a brief literature review of traditional agriculture and management of plant disease/infections. Traditional practices have intense results on advanced agriculture and most of the best practices are based on the traditional way of agriculture. Conservation and resource management have not always been followed using the traditional farming system. The labor-based method is followed in the traditional farming system; an increased number of laborers are involved in agriculture, reducing the chances of unemployment. In [9], the authors discussed the cultivation of fodder crops using the traditional method and partial farming system. The amount of energy consumed by land preparation, fertilizer application, sowing, irrigation, and harvesting is calculated based on ratio. The energy ratio for the traditional and partial method in yield, production energy, input energy, and output energy was calculated. The results indicated that the partial farming system has higher yields compared with the traditional system.

As indicated by the Indian National Crime Record Bureau [10], 60,670 farmers committed suicide in Maharashtra between 1995 and 2013, most of whom were from the cotton-farming belt. The high occurrence of farmer suicides has brought notice to this area and its powerful rural framework. A few families have reacted to the agrarian emergency in Vidarbha by embracing low-input, conventional, and economical farming practices dependent on the standards of sustainable farming, referred to locally as shashwat sheti. While cotton was developed [11] in Vidarbha before colonial rule, its broad development started under pressure from the British organization in an offer to increase agrarian production in India. Chief among the cotton-developing regions was Berar Province, which went under the British Raj in 1853, presently known as the Amravati area of Vidarbha. Cotton from Berar was shipped to the cotton plants of Britain during the Industrial Revolution. Indian cotton production in Vidarbha expanded significantly during the cotton famine that occurred because of the American Civil War between 1861 and 1865.

The Himalayan locale [12] is renowned for its beautiful valleys and quiet magnificence. Patio cultivating provides sustenance for Himalayan individuals, who grow an assortment of oats, millets, beats, oilseeds, and pseudo-grains. A significant part of the varieties grown in the Himalayan zones includes landraces adjusted to the natural local conditions. Additionally, different indigenous practices utilize the available resources. Such practice incorporates field planning, seed planting, weed destruction strategies, determination of yields according to the appropriateness of the ecological conditions, pest control, harvesting techniques, and germination capacity of the seeds. Indigenous individuals, as a rule, develop different crops in an intermixed or interchange way, with the goal that the soil well-being and yield efficiency can be maintained.

According to the 2011 census, there were 1034 ghost villages in Uttarakhand, and that number increased to 1768 in 2018. In the past 10 years, almost 138 individuals moved away from the state's villages. Among these, 33 stay away from their village for the indefinite future and 105 return occasionally to their villages for short spans

[13]. It is not only the relocation of individuals but also their customary information is likewise relocating alongside them. In the recently settled zones, individuals do not practice their customary information and practices. Like this, a wealth of customary knowledge is being forgotten. In the past, farming was the primary wellspring of employment for these individuals, but now farming is considered as a below average occupation. Beforehand, developing was additionally an agreeable endeavor of the residents. Individuals used to take an interest in furrowing and other agriculture acts of each other. In the past, the entire network took an interest to experience the issue of human–untamed life strife.

The interest in freshwater [14] is on the rise because of rapid population growth. Simultaneously, the impact of dangerous atmospheric deviations and environmental changes causes extreme danger to water use and food security. Therefore, irrigation systems are used by numerous farmers worldwide, especially for their detailed measure of water utilization from different sources. There is an expanded spotlight on improving the productivity of water utilization in agriculture water systems. The IoT and propelled control methodologies are being utilized to accomplish improved checking and control of the water system for farming. In this audit, a careful review of water system observations and propelled control frameworks featuring the research of the previous 10 years are introduced. Consideration is paid to ongoing research areas identified from the observations and advance control ideas for a precision water system. Typically, this audit serves as a valuable reference to inform about checking and propelled control openings identified when using water systems in farming and it helps analysts to distinguish directions and gaps for future research in this field.

According to the Economic Survey [15], 2017–18, over half of the entire workforce in India is utilized by the Agricultural division, contributing approximately 17–18% of the nation's GDP. Regardless of the motorization in the agriculture division, improved rural strategies to enhance crop production are needed. Ill-advised water system practices bring about lower food production and imperfect crop yields. Likewise, there is a need to enable the farmers to decide the most reasonable return depending on the plant. One approach to address these issues is to utilize the IoT in the agrarian area. The Internet of Things is the systems administration of sensors, organized features, and gadgets to computerize regular assignments and perform activities contingent upon the condition of the encompassing. The aim of this work is to survey the advancements in agribusiness to computerize the water system and select the best crop to plant.

6.2.2 *Traditional to Smart Agriculture System*

The natural [16] and monetary effects of a few supply frameworks for water systems have been examined. An off-grid (diesel motor) and on-grid (network power) vital supply framework with a PV plant was compared. The LCA technique and restitution period was utilized to decide the effects connected to every supply alternative. An affectability investigation demonstrated the life expectancy length and diesel and power costs impact. The PV framework introduced the least natural effect and life cycle cost, in spite of the fact that it demonstrated the most noteworthy initial speculation.

The communication advances of the IoT assume a significant job in a smart agriculture framework [17]. A smart water system framework dependent on LoRa innovation was proposed and tests were conducted to prove the superb performance of the proposed water system framework. The test results validate the pertinence of the proposed framework and indicate the benefits of LoRa innovation embraced in smart irrigation systems. The framework proposed encourages increasingly active communication, to limit the expense of sending and systems of support. As per the test results, the water system hub outfitted with a hydroelectric generator can work for a considerable length of time.

Regardless of more outstanding efficiency [18], the farming business faces new difficulties that undermine human development. With population growth and the intricacy of environmental change, the farming business has been compelled to progress away from mechanical techniques to information-driven administration and computerization to produce more food while utilizing fewer resources. Another worldview is conceivable with the selection of utilizations and arrangements driven by the convergence of a few key innovations, including the Internet of Things, Artificial Intelligence, and applying autonomy. Together these advancements keep a homestead beneficial and gainful by gathering and breaking down information to assist farmers with dealing with their resources, produce better crop yields and animals while enhancing vitality and compound use and alleviating hazards. Be that as it may, information-driven administration has just been the start of the new change in perspective. The ecological and financial cost of work and resources will continue developing operational proficiency. Farms are moving from information-driven administration to mechanization. The move has been in progress with new applications requiring less human intercession and addressing fundamental issues, including food deserts, work deficiencies, and specialized difficulties, for example, developing in urban conditions and perceiving plants through image processing.

Researchers in [19] indicated that a reason to use smart farming system was because a farmer needs to go to the homestead to check the water level in the field and to turn the water siphon on or off, even at mid-night. This issue can be overcome by improving old farming techniques. Another framework can be created or planned, which changes the old conventional growing into smart development. Researchers attempted to make a straightforward water siphon controller, with the soil moisture sensor, utilizing Esp8266 NodeMCU-12E, which is helpful in the agribusiness field. Framework security was provided by using transport layer security (TLS) and secure attachment layer (SSL) cryptographic conventions. Esp8266 NodeMCU-12E has easy, low force utilization, and a small size microcontroller, which makes the proposed framework suitable for the given application. The high accuracy soil moisture sensor provides simple readings, so soil moisture content is effectively measured. Finally, it showed soil moisture amounts and water siphon state in web page and mobile applications.

Dangers brought about by environmental change [20], growing population, and diminishment of natural resources present a need to search for alternatives that improve manageable crop production. Bioprospecting new assortments that are flexible to the content of the atmosphere and tolerant to biotic and abiotic stresses is of great importance. Supplanting old varieties and finding promising unique assortments could bring a new period of reasonable production. In the national seed framework, especially in the farming world, the assortments utilized by the farmers ought not to be more seasoned than 10 years aside from for certain exceptional circumstances. Furthermore, administrative bodies, breeding associations, seed organizations, and national seed frameworks should be responsible for moving the hereditary additions related to the varietal substitution rate. In addition, a five-point system's viable usage could positively upgrade the VRR, which will eventually lead to reasonable efficiency.

Researchers in [21] explained their challenges in seed priming in field crops. A portion of the difficulties regarding seed adaption preparation involves the life span of seeds after traditional sorts of preparing under surrounding stockpiling conditions and an absence of studies on hermetic bundling materials for broadened capacity. This investigation [22] affirms that steady information about homesteads prompts ideal choices. Rural administration frameworks can deal with farm information so that outcomes are arranged to address revised answers for each farm. This guide for farmers as advanced arrangements joins powers with mechanical technology and artificial intelligence to assist in agriculture. Following 30 years of extraordinary hopes—and dissatisfactions—of applying autonomy to agribusiness, the planning appears to be nearing an end. Nonetheless, to get the most out of Agriculture 5.0, profound preparation should be conveyed to clients; in a perfect world, youthful farmers will be anxious to learn and apply present-day advances to agriculture and conceding a generational restoration still to come. Now is the correct opportunity to push ahead toward advanced and practical agriculture that is fit for demonstrating the full intensity of information-driven administration to confront the challenges presented to food production in the twenty-first century. The advancement to Agriculture 5.0 is in the plan of most significant homestead gear producers for the following decade, and subsequently off-street hardware makers will assume a vital job in this move if agriculture robots are considered as the coming—smarter—age of farm machines. Smart farming technologies raise moral issues related to the expanded corporatization and industrialization of the agriculture area. Specialists investigate the idea of biomimicry to conceptualize smart farming technologies as biological developments that are installed in and as per the indigenous habitat. Such a biomimetic approach of smart farming technologies takes a bit of leeway of its capability to relieve environmental change, while at the same time keeping away from the moral issues identified with the industrialization of the rural area. Six standards of the idea of biomimicry were investigated and these standards were applied with regard to smart farming technologies in [23].

Large data applications in smart farming and recognition of the related financial difficulties to be tended were discussed in [24]. Following an organized methodology, an applied system for examination was built that can likewise be utilized for future investigations on this theme. The survey shows that the extent of big data applications in smart farming goes past essential production; it affects the whole food supply chain. Substantial information is employed to give prescient bits of knowledge in farming tasks, drive continuous operational choices, and update business forms for game-changing plans of action. A few creators in this way propose big data will cause significant movements in jobs and force relations among various players in current food supply chain systems. The partners' scene displays a fascinating matchup between amazing tech organizations, financial speculators, and frequently small new companies and new contestants. Simultaneously, there are a few open foundations that distribute transparent information under the condition that people's protection must be ensured. The fate of smart farming may disentangle in a continuum of two extraordinary situations: shut, exclusive frameworks in which the farmer is a piece of a profoundly coordinated food supply chain or open, cooperative structures in which the farmer and each other partner in the chain organization is adaptable in picking colleagues for the innovation concerning the food production side. The further improvement of information and application framework stages and principles and their institutional implementation will assume a pivotal role in the fight between these situations.

Farming is a tremendous undertaking for humans that affects the lives of the general public [25]. Agriculture is the most significant aspect of social progress. The mechanical headway in remote communication and decrease in size of sensors has extended their use in different fields, such as ecological observing, precision farming, social insurance, military, smart home, and so forth. This work [25] gives an understanding into the different requirements of remote sensor technologies, wireless sensor bits utilized in agriculture, and the difficulties associated with the arrangement of wireless sensor networks (WSN). Smart farming (SF) has assumed a significant role to improve production in agribusiness. This effort centers on smart farming as well as contrasted and customary techniques in agribusiness.

These days, the expanded rural production at a lower cost is increasingly determined by the IoT and the distributed computing ideal models. Numerous explorations and ventures have been explained so far in this unique circumstance. It targets decreasing human endeavors such as resources and force utilization. For the most part, such experiments are dependent on gathering different information relating to the rural territory and sending it to the cloud for additional investigation. Be that as it may, the significant distance between sensors/actuators and the cloud prompts a vast increase in inertness, which prompts a lessening of performance of the irrigation systems, pesticide checking, and so on. This work [26] presents an elective arrangement dependent on fog nodes and LoRa innovation to advance the quantity of hubs organization in smart homesteads. The proposed arrangement decreases the absolute inertness prompted during information transmission toward the cloud for handling.

Smart farming [27] includes the joining of data and communication advances into hardware, gear, and sensors for use in rural production frameworks. Innovations, for example, the IoT and distributed computing, are required to propel this turn of events, presenting artificial intelligence into farming. Accordingly, the points of this paper are twofold: First, to portray the logical information about SF that is accessible in the overall literature dependent on the fundamental variables of advancement by nation and after some time and second, to depict current SF possibilities in Brazil from the viewpoint of specialists in this field. The exploration included directing semi-organized meetings with market and analyst specialists in Brazil and utilizing a bibliometric overview using methods for information mining programming. A combination of the diverse, accessible frameworks available was distinguished as one of the principles restricting elements to SF development. Another restricting variable is the instruction, capacity, and abilities of farmers to comprehend and deal with SF devices. These confinements uncovered an open market door for undertakings to investigate and help tackle these issues, and science can add to this procedure. China, the United States, South Korea, Germany, and Japan contribute the biggest number of research studies to the field. Nations that put more in R&D produce the most distributions; this could show which nations will be pioneers in smart farming. The utilization of both examination strategies in a reciprocal way permitted seeing how science outlined the SF and the main boundaries to embrace it in Brazil.

Since sustainability is challenging and there is a need to secure the production of high-quality, affordable, and healthy food, alternative food production/distribution schemes have emerged that can increase food production without burdening the environment by using technological or organizational innovation [28]. Short food supply chains (SFSCs) and smart farming have potential as solutions for these challenges. Theoretically, introducing smart farming technologies into SFSCs could increase the value-generating capacity of short food supply schemes. However, it is questionable if such technologies are compatible with SFSCs. In this study, the authors followed a mixed research design, to analyze Greek farmers' and consumers' perceptions of the compatibility between smart technologies and SFSCs, and they examined the extent to which compatibility perception affects willingness to engage in smart SFSCs. The authors found that perceived (in)compatibility was central in predicting willingness for both farmers and consumers. The study revealed the existence of two different types of compatibility: actual compatibility, which involves the consistency of smart technologies with the technological advancement of farms and the real everyday needs of farmers, and symbolic compatibility, which involves the meanings attributed to both SFSCs and smart technologies by farmers and consumers. The study concludes that smart technologies are viewed as tools that can lead to a conventionalization of SFSCs, which alters their preferred distinct nature, and that the promotion of smart farming should include more than just traditional views of smart technologies as tools that increase farm efficiency and also pay attention to their compatibility with different types of agriculture and to the ways they can bring about change to farming systems.

Drones are beginning to be used in farming. It is assessed that drones in the rural market will be valued at billions of dollars in the next few years. As editor of the UN Food and Agriculture Organization and the International Telecommunication Union's exploration report on "UAVs and agriculture," data master Gerard Sylvester said that as farmers work to adjust to environmental change and address different difficulties, drones are required to enable all agricultural endeavors to improve proficiency [29]. With the progression in drone advancements and new companies/producers demonstrating interest in the business, it is expected that the expense of the robots and the extra gear will decrease. Likewise, flight time and distance should be increased as a result of basic advancements, for example, battery stockpiling and reduction in payloads. These improvements will guarantee that farmers get more advantages from the utilization of robots in farming [30].

6.3 Supply Chain Management in Agriculture

The study in [31] portrays the overarching promoting courses of action in Tanzania at local, provincial, national, and fare markets utilizing contextual analysis models. The significant obstacles for trade in Tanzania have been sorted into three groups: (1) physical foundation, (2) know-how and capital, and (3) institutional system. A lack of physical foundation including streets increases the expense of transportation, acts as a casual market barrier, creates a wedge between the provider cost and purchaser cost, and expands the loss of short-lived items. The absence of expertise appears in poor market direction and business abilities and prompts challenges in overseeing and getting financial advances. Moreover, the current institutional structure cannot bolster the arrangement of solid brokers and makers' affiliations and other agent bodies to upgrade the limit building and to expect more pleasant terms of exchange. Also, the absence of market data and the weak legal framework lead to challenges in arranging exchange understandings and upholding the current agreements. At present, the important institutional structure has been filled in for by long supply chains of go-betweens and depends on close to home connections between makers, merchants, and intermediaries. To understand the maximum capacity of agrarian exchange as an instrument in the battle against poverty, the recommended approach mediations are to organize and expand financing for a physical framework from both national and universal sources; place accentuation on rural non-farm work and intra-territorial exchange advancement (horizontal integration); advance large-scale limit working in business abilities and market direction; improve access to credit and improve the administration of the current plans; implement the current laws and bolster formalization of agreements to decrease trade dangers; lastly improve dispersion of market data to permit markets to work proficiently.

Supply chain management [32] generally indicates dealing with the connection between organizations liable for the proficient production and supply of agribusiness items from farm level to buyers, to dependably meet buyers' necessities as far as amount, quality, and cost. This regularly incorporates the administration of both horizontal and vertical coalitions. In developing nations, the supply chain of rural items regularly includes numerous players or operators with numerous farmers on one side and shoppers on the other. These conventional supply chains are firmly connected with social structures. For the most part, small farmers in developing nations are not able to set their own price and their contact with "business sectors" is regularly restricted to associating with a produce gatherer or to deals at the neighborhood/town market and area advertising. Indian agribusiness is ruled by small holders, approximately 86% having the land property of up to 2 ha, with a normal, size of the land property of 0.53 ha according to the Government of India. The survivability of these small holders is questionable because their failure to access markets is a significant restriction. The broadening of farming toward high value wares has been proposed as a conceivable choice for small holders. A move is occurring in the types of food Indian buyers prefer, including a nutritious eating regimen of organic products, vegetables, milk, and meat. Continued financial development and a quickly developing urban populace are powering fast development for sought after high worth food items. Moreover, the exported portion of horticulture and animal food products in agriculture products expanded from 24% in 1981 to 35% in 2003. The farmer's portion of the wholesale price continues to only be approximately 35%, with the significant portion going to middlemen as promoting cost, as a result of wasteful supply chains. Thus, supply chain management might be an incredible asset in connecting farmers to the business sectors to increase their pay.

Third-party logistics (3PL) specialist organizations [33] can assume a significant job in the agriculture supply chain management for consumer loyalty and cost reduction in overseeing supply chains. Determination of the providers is one of the significant elements that should be thought of because choosing the best 3PL providers can expand the competitiveness and sustainability of the supply chain. In any case, this choice becomes confused when there are numerous 3PL providers having different models and incorrect boundaries. Also, the vagueness and suspiciousness of the experts' feelings exacerbates the issue. In this way, a dynamic device dependent on multi measures has been utilized generally as a fuzzy logic, progressive procedure approach to determine the best 3PL specialist co-ops. The principle point of this exploration is to build up a blueprint of various rules for provider's choice dependent on literature reviews and methods utilized to choose the best 3PL providers. Moreover, this study gives an increasingly exact, compelling, and proficient decision help tool for choosing the best 3PL suppliers. This exploration may help in expanding the inclination to outsource logistics activities to upgrade the manageability of the IoT-based agribusiness supply chain.

The authors in [34] investigated and analyzed factors influencing cauliflower post-harvest losses (PHL) at the farm level in the Surat area of Gujarat. Sample data was gathered from 120 cauliflower producers from 12 towns using the survey method. Multiple regression analysis was carried out to decide the indicators of factors that lead to fresh produce PHL. The main reasons for PHL at maker level were: damage during harvest, damage because of disease, damage because of pests, damage during transportation; and an absence of legitimate cleaning and washing indicated a positive connection between every autonomous variable and the reliant variable. To diminish PHL, farmers need to focus on reducing the above concerns because these factors were shown to have a noteworthy impact on the amount of PHL of fresh produce.

Blockchain technology (BT) has affected the supply chain by removing the trust-related issues [35]. Studies are being performed worldwide to use the advantages BT can provide to improve the presentation of the supply chains. The literature has shown that BT offers different benefits prompting enhancements in the sustainable performance of the agribusiness supply chains. Usually, BT will bring a change in perspective on how the transactions are conveyed in the agriculture supply chains (ASC) by decreasing the high number of middlemen, deferred installments, and high transaction lead times. India, a developing economy, must consider the food security needs of an ever-developing populace and address numerous difficulties influencing ASC supportability. Therefore, it is necessary to embrace BT in the ASC to use the different advantages. The authors recognized and built up the connections between the empowering agents of BT selection in ASC and identified 13 empowering influences from the literature, which were approved by the specialists before applying a joined interpretive structural modeling (ISM) and decision-making trial and evaluation laboratory (DEMATEL) procedure to imagine the complex causal connections between the distinguished BT empowering influences. The authors found that, among the recognized empowering agents, traceability was the most critical purpose behind BT execution in ASC, followed by auditability, immutability, and provenance. The discoveries of the investigation will assist the experts in designing BT execution methodologies in agribusiness, making an ongoing information-driven ASC. The outcomes will likewise help the policymakers create arrangements for quicker usage of BT guaranteeing sanitation and practical ASCs.

Agriculture supply chain management involves everything associated with the process of moving farm products from the field to the client, and it is an important part of a nation's economy [36]. This study aimed to determine the difficulties present in the farming supply chain in India based on reviewing the literature and the Delphi method. The authors then used the decision-making trial and evaluation laboratory approach to show the determined difficulties, investigate the cause–effect interrelationship, and to build up the deliberate progressive structures of difficulties through an interpretive structural modeling strategy. Considering the Indian setting, two elements, specifically limited mixing among the national agriculture markets and restricted farming market infrastructure, were determined to be the most significant ones. The incorporated model acquired as a yield of this examination plans to direct the agrarian arrangement and chiefs to improve the presentation of the rural supply chain in India. Additionally, some basic proposals have been given to improve the proficiency of the rural supply chain management.

6.3.1 Farmer to Factory (F2F)

The agriculture sector is made up mostly of family businesses, which face difficulties participating in worldwide business because of the Free Trade Agreement (FTA) [37]. In this paper, the authors suggest business strategies to help the family businesses compete in the worldwide value chain. The researchers conducted a literature review on family-owned businesses and farming to determine the challenges they faced and propose systems that can help family farms compete. Two main challenges were determined, in particular the use of labor and capital. It was found that Indonesia's family-owned businesses used a higher amount of labor than worldwide businesses, on average. Even so, the all-out estimation of capital use in Indonesia's family businesses was lower than worldwide businesses because farmers cooperated and shared know-how. Capital use was found to be driven by social capital fortification, capital allotment, and capital accessibility, and the Indonesian family-owned businesses were found to not be able to take advantage of capital use. The authors proposed a procedure for the family-run company to defeat the danger of work through its capital utilization quality. Addressing these challenges should help to fortify family-owned businesses in the agriculture sector to conquer worldwide trade challenges and engage in new opportunities. This work may benefit governments in drawing up arrangements for helping the family-owned businesses in the farming sector to become global. This work adds to the scarce literature related to family farming businesses in Indonesia.

The sustainable development goals (SDGs) [38] are a significant system that sets the development plan for the following 15 years. Although the SDGs seem to address improvement challenges in a thorough manner, they have many flaws. This is clear in their origination and comprehension of sustainability, particularly according to agribusiness and its duty to reduce poverty and promote development. By concentrating on the agriculture segment of the SDGs, the author shows how the interests of agribusiness are expressly adjusted and progressed through a flawed view of sustainability that is found in the SDGs. The author suggested that the "agriculture for development" plan available in the SDGs is more about guaranteeing the interests of agribusiness to the detriment of guaranteeing genuine sustainable development. The author concludes that the social and political endeavors originating from such an arrangement sabotage endeavors to progress genuine sustainability and meaningful ecological relations.

This year the fund subject gathered information on extra regions that are fundamental to agrarian accounts; however, global procedures are not completely evolved. Partial credit guarantee frameworks and rural loaning shares are two zones of the account point examined. Partial Credit Guarantees (PCGs) can be a useful asset to build credit for agribusiness. They diminish the hazard that monetary establishments take when loaning to farmers and agribusinesses by going about as an insurance substitute, wherein "if the borrower neglects to reimburse, the moneylender can fall back on partial reimbursement from the underwriter." However, the presence of a PCG does not ensure expanded farming division loaning; rather, PCG plan and usage effectively affect program manageability and adequacy. Because there is no "one-size-fits-all" structure for PCGs, the group decided not to score this information. The information gathered shows that 18 of the 62 nations reviewed have a PCG specifically for rural advances given by business banks. Only two high-income nations, Italy and Korea, have PCGs.

The reason to make an agriculture marketing organization is to help overcome any barrier among producers and buyers. The general murkiness surrounding production has created a rift between the industry and consumers, leading them to scrutinize their food, and a general lack of understanding about the farming lifestyle. A rural promoting firm in California is attempting to close the gap between producer and consumer by speaking to and advocating for local businesses. In general, the American farmer exceeds expectations at numerous things; however, they often fall short regarding communication. Current farmers have had the option to create more food than any time in recent memory in spite of fewer resources and inputs. As the population grows and more individuals relocate from rural to urban areas, the populace engaged with agribusiness declines, and a knowledge gap is made. Essentially because of where individuals live, there is a distinction between general society and the story of farming. The average farmer is seen as a blockhead, who does nothing but farm and sleep [39]. Today, farming proceeds as a generational business based upon previous family achievement and having a gainful future for future generations — farming remains a lifestyle. Agriculture faces a purported farm issue. A farm issue is a barrier to better open comprehension by the urban majority of the country of the issues and needs of agribusiness.

This study aimed to distinguish a connection between agrarian added value share and the agriculture employment share [40] through a correlation analysis utilizing the Pearson's correlation coefficient. The results indicate that the lower the portion of the agriculture added value in the Ecuadorean economy, the greater the portion of the employment in the area, which implies that Ecuadorian farm workers generally become poorer relative to others laborers over time.

The success of some urban farmers has attracted worldwide intrigue. The authors conjectured that cultural inclinations and the worthiness of urban farming undertakings and items decide the achievement or disappointment of urban agriculture businesses. The authors studied 386 urban participants in Berlin, Germany, to determine general inclinations for the gainful utilization of urban space, the acknowledgment of various urban agriculture structures, and requests and desires with respect to agriculture items. The outcomes show that more than 80% of the respondents favored having open frameworks, for example, open green spaces, intercultural nurseries, and rooftop gardens. In fact, land uses that do not give openness, for example, glades, aquaponic farms, or concentrated rural and green scenes were less favored (under 40%). While 60% of members communicated acceptance of rooftop gardening, agribusiness in the urban periphery, or in downtown brownfields, 65% dismissed having farming in multi-story structures, agro parks, or aquaponic farms [41].

Agriculture assumes a critical job in Metropolis Ruhr cultivating approximately 33% of the metropolitan zone; however, on-going loss of farmland and transient rent of land influence farms extensively by complicating access to land. Generally separated and differentiated farms require a certain measure of farmland to be effective, which is compromised by further farmland losses in addition to expanded rivalry for the rest of the farmland [42]. Long haul arranging security is pivotal for farms that focus on high added value crops and organic farming. Dynamically developing participatory farming activities exhibit urban farms' capacities to inventively adjust to cultural preferences. Better knowledge of proficient urban agribusiness' homestead exercises and plans of action is of significance for farms and their related advisory services to more accurately address the urban setting in farm advancement methodologies, in addition to aiding public authorities' land-related decision-making with regard to planning and polices. The economic viability of expert urban farming is the key prerequisite for extra social, environmental, and landscape improvements inside urban territories.

6.4 IoT-Based Smart Farming

The flowchart of the IoT-based farming system is graphically represented in Fig. 6.2. Run of the mill business sensors for agribusiness irrigation system frameworks are extremely expensive, making it inconceivable for small farmers to utilize them [43]. Although, producers are currently offering low-cost sensors that can be associated with hubs to implement affordable frameworks for irrigation management and to monitor farms. In light of the ongoing advances in IoT and WSN innovations applied in the improvement of these frameworks, the authors presented an overview of the

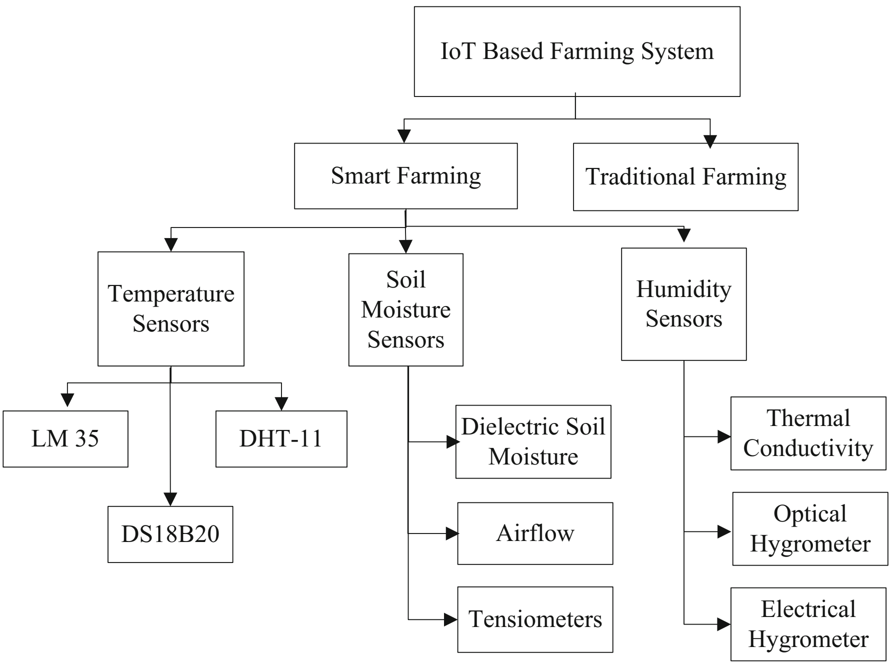


Fig. 6.2 IoT-based farming system

current best-in-class smart irrigation systems. The parameters that are monitored in irrigation systems include water amount and quality, soil attributes, and climate conditions. The researchers give an overview of the most used hubs and wireless technologies and examine the difficulties and the accepted procedures for the usage of sensor-based irrigation systems.

6.4.1 Arduino UNO

Arduino UNO is an Atmega328-based microcontroller board created by Arduino.cc. Sensors and different gadgets are incorporated into the Arduino adding to its incredible on-board capabilities. It has 6 simple pins and 14 digital pins. Arduino IDE is utilized to program Arduino Uno. Its working voltage is 5 volts. Arduino UNO reads the qualities from the sensors, and dependent on these qualities, it will train the engine driver to turn on/off the engine siphon. Figure 6.3 shows the IoT circuit in agriculture.

A soil moisture sensor, humidity sensor, and water level sensor are connected with the Arduino UNO board through wires and breadboard. The soil moisture sensor will detect the moisture of the soil. A limit has been set for both least and most extreme so that at whatever point the moisture content crosses the predefined threshold limit, the engine will be turned on/off. The water level sensor functions in the same way, where a base and highest limit have been set so that at whatever point the water level crosses the predefined threshold limit, the engine will be turned on/off to fill the tank. An LCD is associated with Arduino and all the sensors to show the status of moisture content in humidity, soil, and water level in the tank. A 5 V engine siphon was utilized for this work because this is a structure model, and the Arduino used in this work can give a limit of 5 V power. Figure 6.4 illustrates the schematic design of a smart farming prediction system.

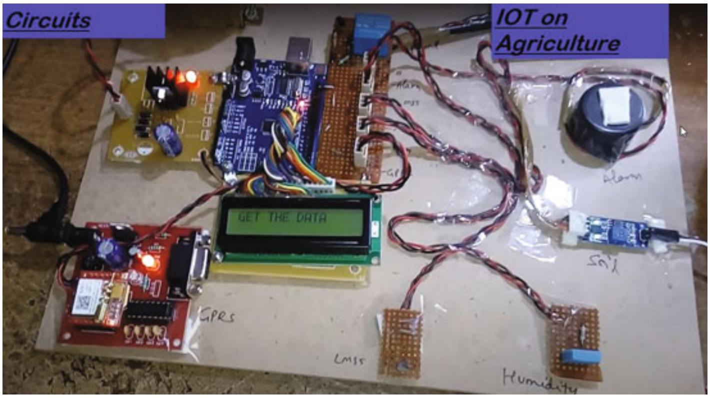


Fig. 6.3 IoT in agriculture

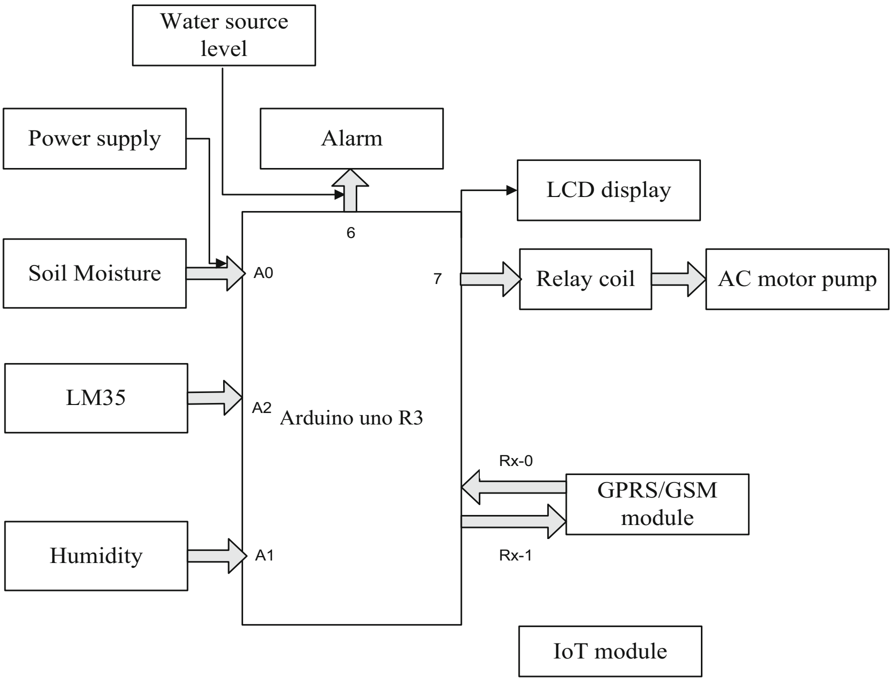


Fig. 6.4 Schematic design of smart farming prediction system

Figure 6.5 explains the workflow of the proposed smart farming prediction system.

Table 6.1 gives a summary of various sensors used in smart farming. LM35 is a temperature sensor that yields a simple sign that corresponds to the immediate temperature. The yield voltage can acquire a temperature reading in Celsius. The benefit of LM35 over thermistor is that it does not require any outer alignment. LM35 temperature sensor has 4-20 V, OUT, and GND. For the working principle of DS18B20 temperature sensor, the default goal at power-up is 12-piece. The DS18B20 controls up in a low force inactive state. Following the change, the following warm information is put away in the 2-byte temperature register in the scratchpad memory, and the DS18B20 returns to its inactive state. The moisture-detecting part of the DHT11 is a moisture holding substrate with the terminals applied to the surface. The adjustment in the opposition between the two terminals corresponds to the relative humidity. The reading for a dry soil is somewhere in the range of 3 and 5, approximately one for air, and approximately 80 for water. The most reasonable alternative for estimating soil moisture content is robust state sensors, which cost approximately $35–60 for each sensor unit. Different kinds of soil properties, including compaction, structure, soil type, and moisture level, produce interesting distinguishing marks. At the point, when covered in the dirt, the clay tip of the tensiometer permits water to move openly in or out of the cylinder.

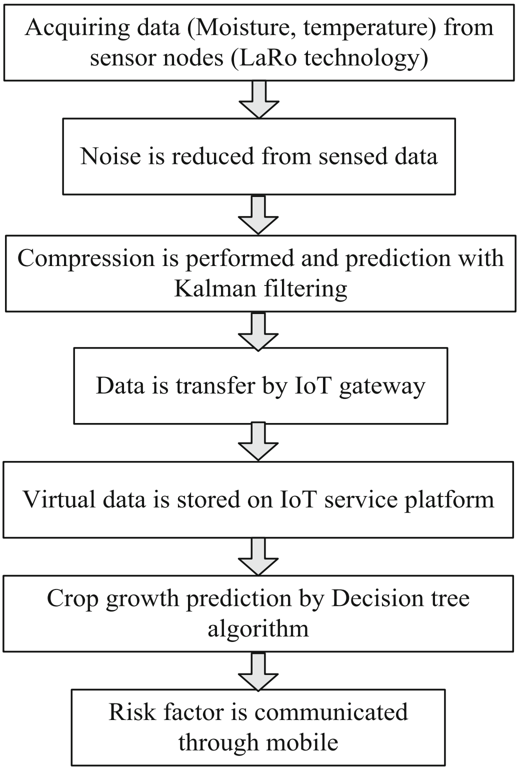


Fig. 6.5 Workflow of proposed system

As the soil dries out, water is sucked out through the porous clay tip, making a partial vacuum inside the tensiometer, which is read on the vacuum measure. In the thermal conductivity humidity sensor, one thermistor is hermetically fixed in a chamber loaded up with dry nitrogen while the other is presented to open condition through little vent openings. An optical hygrometer gauges the retention of light by water from the surroundings. A light producer and a light locator are orchestrated with a volume of air between them. In an electrical hygrometer, generally, a sensor estimates changes in a layer of lithium chloride or another sort of semiconductor.

6.5 Results

The hardware setup of the proposed smart farming prediction system is shown in Fig. 6.4. An Arduino board set up is used for an automatic or manual function to predict crop growth. The prediction is based on sensors for temperature, soil moisture, and humidity. Water source-level data is fixed before the buzzer so that the information is transferred according to the water level. 16x2 LCD is used to display farming risk factors, which are sent through a mobile device. Relay coil and Ac

Table 6.1 Summary of sensors used in agriculture

Sensors	Sensor Types	Sensors	Functions
Temperature	LM35		It can measure temperature more precisely than utilizing a thermistor
	DS18D20		Reads with a precision of ±0.5 °C from –10 °C to +85 °C and ± 2 °C exactness from –55 °C to +125 °C.
	DHT-11		It is an essential, ultra-minimal effort computerized temperature and humidity sensor
Soil moisture	Dielectric soil moisture		It works by utilizing two cathodes to quantify the electrical opposition in the soil
	Airflow		It measure soil air penetrability
	Tensiometer		Uses probes to measure soil compaction
Humidity	Thermal conductivity		Measure the warm conductivity of both dry air and air with water vapor
	Optical hygrometer		The retention of light by water from the surroundings
	Electrical hygrometer		Estimations on changes in electrical obstruction or capacitance

motor pumps are fixed for reliability. In an IoT module, GSM.GPRS is invoked for authentic communication. Ranges for automatic functions of temperature increased by more than 70, communication is sent through a mobile device. If the humidity range is more than 50 or the soil moisture range is more than 160, then this will be communicated. Based on the threshold value of each sensor, a farmer can predict water source level for various crops. In summer season, more water is required for paddy fields, whereas less water is consumed during winter.

Figure 6.5 explains the workflow of the proposed smart farming prediction system. Data is acquired from sensor nodes using LoRa technology. Noise is reduced from the sensed data. Compression takes place for prediction using Kalman filtering (KF). In this work, nominal packet size is utilized for transmission to accomplish better pressure; however, the Kalman filter sends every second at least, which can lessen the sum of energy needed for every exchange. Rather than utilizing models dependent on schedule for predication, the KF-based procedure utilizes a total model for predication where a few states are obscure, which gives a more detailed prediction investigation. Because of the decrease in the pace of transmission, the reproducibility of information is diminished. To accomplish better quality, KF eliminates noise from detected information gained by leaf hub and makes the reproduced sign more precise compared to unpleasant perceptions. Data is transferred by the IoT gateway. Water source level data is stored on the IoT service platform. A decision tree algorithm is performed to predict crop growth and, finally, the farming risk factor alert is communicated through a mobile device.

The output result of prediction analysis based on maintenance cost per acre along with predicted market price per kilogram in rupees is shown in Fig. 6.6. Various crops are shown in the x-axis such as paddy, wheat, groundnut, corn, and onion. The value of maintenance cost for paddy is 201 and market price is 75. Maintenance cost for wheat is 229 and market value is predicted as 26. For groundnut, 164 is the maintenance cost and 137 is the predicted market value. The value of maintenance cost for corn is 260 and the market price is predicted as 42. The maintenance cost for onion is 257 and the predicted market price is 50. The proposed smart farming prediction system predicted maintenance cost is high and predicted market price is low for corn. The crop corn has a high risk factor according to the proposed sensors.

In the Harvard setup, the water source level data is used to predict the required water for the particular corp. Figure 6.7 shows the maximum water required per day in millimeters. Experimental results show paddy required 1350 mm or maximum water. Wheat required 1267 mm of water, groundnut required 600 mm of water, corn required 430 mm of water, and onion required 126 mm of water. Based on the results, farmers can easily predict which crop is suitable for farming.

A farmer can easily understand and predict the farming system from Table 6.2 based on the suitable season, water required per day, the temperature in Celsius, maintenance cost per acre, a market predicted value of the particular crop, which is mentioned in rupees per kilogram, and predict the risk factors of the particular crop to yield or not. This prediction system is useful to the farmers to prevent loss in production. From the chart, a farmer can predict suitable crops from the soil condition and the water requirements. The irrigation system may by via pond, well,

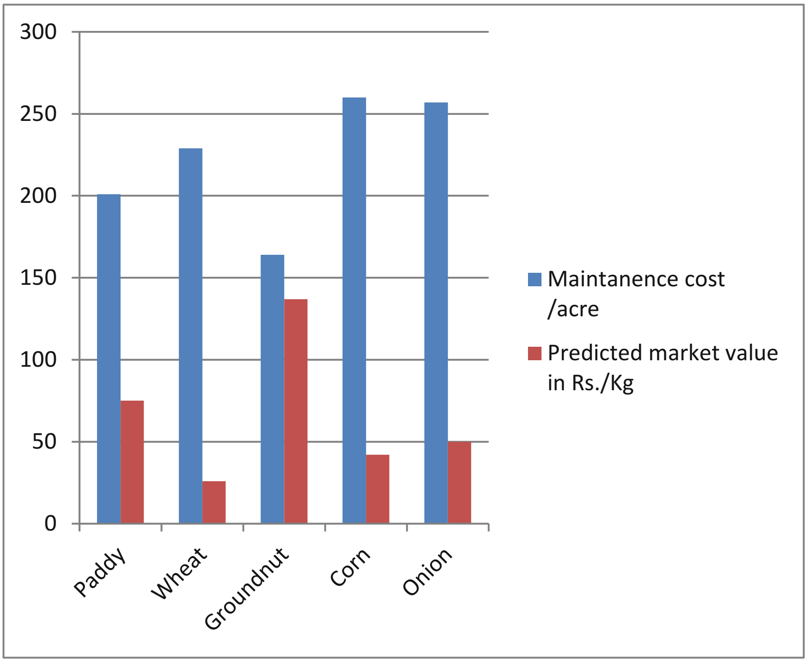


Fig. 6.6 Risk factor of the crop

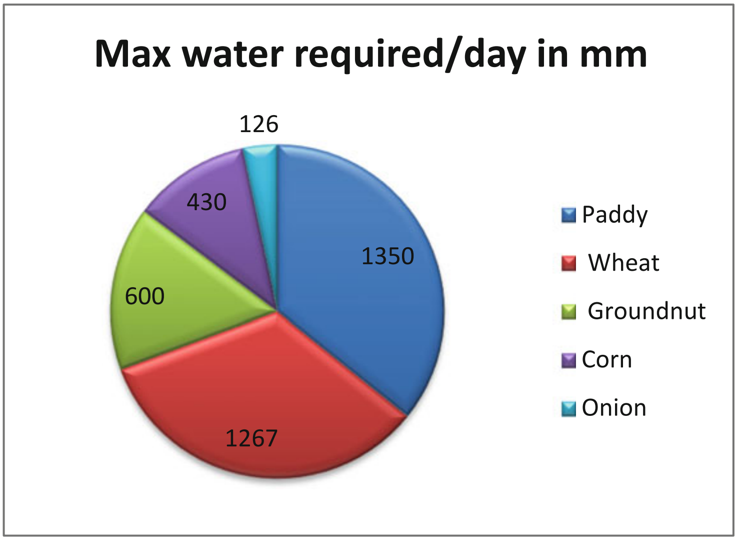


Fig. 6.7 Water consumption based on the crop

Table 6.2 Farmers prediction chart

Crop	Season	Water required/day	Temperature in Celsius	Maintenance cost/acre	Predicted market value in Rs./Kg	Risk factor
Paddy	Summer	1100-1250 mm	20–27	201	75	Low
Wheat	Winter	130.7–136.7 mm	21–26	229	26	Low
Groundnut	Summer	400-600 mm	27–30	164	137	Low
Corn	Spring	2.5–4.3 mm	18–23	260	42	High
Onion	Winter	10-12 mm	12–25	257	50	Moderate

or river according to the land. Agri-IoT-based sensors can be used to sense weather conditions, water status, soil moisture, and humidity.

6.6 Conclusion

In this chapter, we have depicted Agri-IoT, an IoT-based structure applying ongoing stream handling, examination, and thinking in the space of agriculture, in light of sensors, encouraging increasingly educated and exact dynamic by farmers and occasion choice. We have examined the presentation of IoT in smart farming and its chances, through the consistent blend of heterogeneous advances and the semantic mix of data from different sources such as sensors, online life, associated farms, administrative alarms, guidelines and so on, guaranteeing an increase in production and efficiency, better quality items, environmental protection, less utilization of resources(for example, water , fertilizer, and energy), and quicker response to capricious occasions, as well as provide greater transparency to the purchaser.

References

1. A.G. Koutroulis, L.V. Papadimitriou, M.G. Grillakis, I.K. Tsanis, K. Wyser, R.A. Betts, Freshwater vulnerability under high end climate change. A pan-European assessment. Sci. Total Environ. **613**, 271–286 (2018)
2. S. Adamala, Traditional to smart irrigation methods in India: Review. Am. J. Agric. Res. **1**(1), 0013–0023 (2016)
3. M. Katoch, Farming too can be quite rewarding – both mentally and financially. https://www.thebetterindia.com/125477/kisan-diwas-successful-farmers-lucrative-business/. (2017)
4. Mobasshir & Mahbub. A smart farming concept based on smart embedded electronics, internet of things and wireless sensor network. Internet of Things. **9**. (2020)
5. K.P. Chandan, & B. Roheet, Social internet of things in agriculture: An overview and future scope. Studies in Computational Intelligence. pp. 317–334. (2019)
6. M. Swain, R. Singh, A.K. Thakur, A. Gehlot, A machine learning approach of data mining in agriculture 4.0. Intern. J. Emerg. Technol. **11**(1), 257–261 (2020)
7. D. Welle, (n.d). Antarctic greenhouse vegetables picked. Retrieved 1 Dec 2019, from https://www.dw.com/en/scientists-harvest-antarctic-greenhousevegetables/a-43265721
8. Thurston, H.D. Sustainable practices for plant disease management in traditional farming systems. Book Chapter. (2019)
9. P.K. Mishra, S. Sharma, H. Tripathi, D. Pandey, Energy input for fodder crop productions under different types of farming system. Plant Arch. **19**(1), 1358–1362
10. P.B. Behere, 'Farmers' Suicide in Vidarbha Region of Maharashtra State: A Myth or Reality? Indian J. Psychiatry. **50**(2), 124–127 (2008)
11. B.B. Mohanty, We are like the living dead: Farmer suicides in Maharashtra, Western India'. J. Peasant Stud. **32**(2), 243–276 (2005)
12. A.K. Singh, R. Rana, Nationally important agricultural heritage systems in India: need for characterization and scientific validation. Proc. Indian Natl. Sci. Acad. **85**(1), 229–246 (2019)
13. K. Singh People migrating from villages each day, 734 new 'ghost' villages in Uttarakhand. https://timesofindia.indiatimes.com/city/dehradun/138-people-migrating-fromvillages-each-day-734-new-ghost-villages-in-uttarakhand/articleshowprint/64044856.cms. Accessed on 5 May 2018
14. E. A. Abioye, M.S.Z. Abidin, M.S.Z. Mahmud, S. Buyamin, M.H.I. Ishak, M.K.I. Rahman, A. O. Otuoze, P. Onotu, M.S.A. Ramli, A review on monitoring and advanced control strategies for precision irrigation. Computers and Electronics in Agriculture. pp. 1–22. (2020)
15. S. Pandey, A. Shrivastava, R. Vijay, S. Bhandari, A review on smart irrigation and crop prediction system. International Conference on Sustainable Computing in Science, Technology & Management (SUSCOM-2019). Amity University Rajasthan, February 26–28, 2019
16. A.M. Garcia, J. Gallagher, A. McNabola, E.C. Poyato, P.M. Barrios, J.A.R. Diaz, Comparing the environmental and economic impacts of on-or off-grid solar photovoltaics with traditional energy sources for rural irrigation systems. Renew. Energy **140**, 895–904 (2019)
17. W. Zhao, S. Lin, J. Han, R. Xu, L. Hou, Design and implementation of smart irrigation system based on LoRa. IEEE Globecom Workshops. (2017)
18. I. Charania, X. Li, Smart farming: Agriculture's shift from a labor intensive to technology native industry. Internet of Things. Elsevier. 9. pp. 1–15. (2020)
19. R. V. Kodali, B. S. Sarjerao, A low cost smart irrigation system using MQTT protocol. TENSYMP. IEEE. (2017)
20. R.P. Singh, A.D. Chintagunta, D.K. Agarwal, R.S. Kureel, S.P. Jeevan Kumar, Varietal replacement rate: Prospects and challenges for global food security. Global Food Security. pp. 1–7. (2019)
21. M. Farooq, M. Usman, F. Nadeem, H. Rehman, A. Wahid, S.M.A. Basra, K.H.M. Siddique, Seed priming in field crops: Potential benefits, adoption and challenges. Crop Pasture Sci. **70**, 731–771 (2019)

22. V. Saiz-Rubio, F. Rovira-Mas, From smart farming towards agriculture 5.0: A Review on crop data management. Agronomy. (2020)
23. V. Blok, B. Gremmen, Agricultural technologies as living machines: Toward a biomimetic conceptualization of smart farming technologies. Ethics Policy Environ. **21**(2), 246–263 (2018). https://doi.org/10.1080/21550085.2018.1509491
24. S. Wolfert, L. Ge, C. Verdouw, M. Bogaardt, Big data in smart farming – A review. Agr. Syst. **153**, 69–80 (2017)
25. T. Rajasekaran, S Anandamurugan. Challenges and applications of wireless sensor networks in smart farming – A survey. Advances in Big Data and Cloud Computing, Advances in Intelligent Systems and Computing. pp. 353–361. (2019). https://doi.org/10.1007/978-981-13-1882-5_30
26. M. Baghrous, A. Ezzouhairi, N. Benamar, Smart farming system based on fog computing and LoRa technology. Advances in Intelligent Systems and Computing. Proceedings of ESAI 2019, Fez, Morocco **1076**, 217–225 (2019)
27. D. Pivoto, P. Waquil, E. Talamini, C. Finocchio, V. Corte, G. Mores, Scientific development of smart farming technologies and their application in Brazil. Information Processing in Agriculture. pp. 3–35. (2017)
28. E. Lioutas, C. Charatsari, Smart farming and short food supply chains: Are they compatible? Land Use Policy **94** (2020). https://doi.org/10.1016/j.landusepol.2020.104541
29. Q. Ren, R. Zhang, W. Cai, X. Sun, L. Cao. Application and development of new drones in agriculture. IOP Conf. Series: Earth and Environmental Science. 440. (2020)
30. H. Pathak, G.A.K. Kumar, S.D. Mohapatra., B.B. Gaikwad, J. Rane, Use of drones in agriculture: Potentials, Problems and Policy Needs. ICAR-National Institute of Abiotic Stress Management. (2020)
31. E. Eskola, Agricultural marketing and supply chain management in Tanzania – A case study. ESRF Study on Globalization and East Africa Economies. (2005)
32. S. Singh, B.K. Sikka, A. Singh, Supply chain management and indian fresh produce supply chain: opportunities and challenges. International food & agribusiness management association, 19th Annual World Symposium Budapest, Hungary, June 20–23. (2009)
33. S.Y. Garg, S. Luthra, Selection of third-party logistics services for internet of things-based agriculture supply chain management. Intern. J. Logistics Syst. Manag. **35**(2), 204–230 (2020)
34. C.R. Asmitaben, N. Singh, V.M. Thumar, Factor contributing post-harvest losses in supply chain management of Cauliflower in Gujarat. J. Pharmacognosy Phytochem **9**(2), 2275–2277 (2020)
35. S.S. Kamble, A. Gunasekaran, R. Sharma, Modeling the Blockchain enabled traceability in agriculture supply chain. Int. J. Inf. Manag. (2019). https://doi.org/10.1016/j.ijinfomgt.2019.05.023
36. B.B. Gardas, R.D. Raut, N. Cheikhrouhou, B.E. Narkhede, A hybrid decision support system for analyzing challenges of the agricultural supply chain. Sustain. Product. Consumption (2019). https://doi.org/10.1016/j.spc.2018.11.007
37. N. L Tirdasari, D. Indrawan, I. Fahmi, Family business in agriculture: challenge and strategy to face global business. advances in economics, Business and Management Research. 3rd International Conference on Trade (ICOT 2019). Vol. 98. (2019)
38. M. Spann, Politics of poverty: The post-2015 sustainable development goals and the business of agriculture. Globalizations **14**(3), 360–378 (2017). https://doi.org/10.1080/14747731.2017.1286169
39. E. Larson, M. Alameda, Business model: California agriculture marketing agency. Business Plan. (2017)
40. D. J. Pablo, The bad business of agriculture a correlation analysis on employment share and agriculture added value share in Ecuador. Revista Politecnica. **37**(2). 2016
41. K. Specht, T. Weith, K. Swoboda, R. Siebert, Socially acceptable urban agriculture business. Agron. Sustain. Dev. **36**(17) (2016). https://doi.org/10.1007/s13593-016-0355-0

42. B. Polling, M. Mergenthaler, W. Lorleberg, Professional urban agriculture and its characteristic business models in Metropolis Ruhr, Germany. Land Use Policy **58**, 366–379 (2016)
43. L. Garcia, L. Parra, J. M. Jimenez, J. Lloret, P. Lorenz, IoT – based smart irrigation systems: An overview on the recent trends on sensors and IoT systems for irrigation in precision agriculture. Sensors. pp. 1–48. (2020)

Index

S. Paul et al. (eds.), *Frontiers of Data and Knowledge Management for Convergence of ICT, Healthcare, and Telecommunication Services*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-77558-2