Prediction of effort required to design software for smart city applications

P Rai, S Kumar, D K Verma
Jaypee University of Engineering and Technology, India
prerana.rai99@gmail.com, dr.shishir@yahoo.com, dinesh.hpp@gmail.com

Abstract. In present era smart cities are essential for further development. The fast growth in the urban population is a serious concern for city administration. It leads to various challenges while handling problems of resource crunch and managing them. Developments that relate to smart cities are helping the administration address these problems. Its initial effort prediction and its scalability are the main features for the completion of a successful smart-city project. The initiative depends on many factors that are inter-related. It is not possible to change or optimise certain kinds of aspects separately. The last-minute change in the schedule raises the project budget and time. In the early phases of the project, a realistic and concrete project plan needs a fair estimate of effort. This projected effort helps to monitor and handle the activities of application creation by the developer. Unfair allocation of efforts and resources between development activities, including software projects, contributes to the failure of smart city projects, too. Therefore, for smart city ventures, the secret to success is an accurate estimation of implementation activities at the earliest. Numerous approaches are available for effort prediction of software as well as hardware projects used to develop the smart city projects. Hence, the gap in the different methods available creates thrust areas for researchers to investigate and propose an effort prediction model for smart city-based projects. In this paper, a model has been formulated that takes into consideration a set of critical factors for a given IoT based application used to develop smart city in the subject, and it further provides estimation for the effort.

Keywords. Smart city; Traffic congestion; Particle Swarm Optimization; Support Vector Machine; Software Effort Estimation

1. Introduction
A change in the way of treating the surrounding world can be seen in the last twenty years, the perception of the external environment by society has been increased. The rapid growth of manufacturing, transport or services has transferred to daily life and, above all, to quality. As per the upgradation of technology and day to day involvement of it in life. People are getting dependent on technology and are in search of making the daily chores easier and done in smart way[1]. Smart city makes the lifestyle whole lot easier by providing a society which has been totally dependent on technology for every work. This gives a smart and intelligent way of living and working in the city.

A smart city is the integration of information and communication technology on the physical infrastructure. This combination helps in providing an efficient service with a sustainable high living standard for the citizens. A city is considered as “Smart” when the investments in human and social
capital encourage sustainable economic development with a more top quality of life. This high-quality life concern the prudent management of natural resources through participatory government [2]. A smart city taxonomy comes from four different perspectives: application domain, data processing, technical infrastructure, and system integration. These aspects are companied by three distinct layers, i.e., the upper layer consists of providing services which include processes and relevant activities related to societal needs. The middle layer consists of cyber-infrastructure, which consists of all computer and communication techniques to inter-connect with different components. The lower layer consists of physical infrastructures, which includes physical amenities that already exists as a part of an ecosystem in the city [3].

In a smart city, many different projects are carried on different scale. There are several metropolitan cities in the world those have successfully launched different projects for environmental monitoring. The challenge faced to take these project initiatives is long-term sustainability, especially from a pilot phase to large scale deployment. The condition becomes more critical when the external funding is suspected and sometimes a pilot project lose vendors interest when its success is described [4]. In this regard, sustainability evaluated at the early stage of the project with a long-term perspective. The administrator, who approved or disapproved the plan, do not take an interest in an expensive one whereas low-cost projects create more interest. The effort required for developing a project for smart city, plays a very vital role in distinguishing expensive and inexpensive project. Also, predicting the resource utilization at the initial level of plan, helps in making the right decision of whether to invest in it or not [5]. Predicting efforts and resources also provides the success or failure rate of the project with detailed information about the resources, time, and human resources being used to develop the model.

Many researchers [6][7][8] are coming up with different application plans to design the smart cities in a very efficient and intelligent way. Few applications, which helps in reducing the resource, have already been launched in the market. The devices are as follows: for climate -EverImpact [9], Smart Street Lighting – Telensa [10], Luminext [11], Echelon [12], Transportation – BestMile [13], Kiunsys [14], Anagog [15], Smart Parking – Streetline [16], Waste Management – Enevo [17], Bigbelly [18], Compology [19], Waste Water Management –Ayyeka [20]. Multi-national company IBM is also working on a smart city project “SCRIBE,” which defines the smart city in terms of semantic model based on the data gathered from around the world [21]. The ontology of SCRIBE is defined using open standards such as the Common Alerting Protocol and the National Information Exchange Model (NIEM). It addresses the heterogenous data issue of different smart city domains [22].

In this paper, a model for making a smart city application is discussed, which comprises all the essential requirements that need to take into consideration for predicting the efforts for it. It helps in evaluating the effort of software at the initial phase of development. The paper is structured as follows: Section 2 highlights the methodology used to develop the model. Section 3 describes the smart city-based software applications. Section 4 explain the formulation of the proposed model with the detailed explanation of the parameters used for analyzing the work. Section 5 describes the result obtained from the proposed model. Section 6 describes the conclusion extracted from the proposed model.

2. Methodology Used

2.1. Support Vector Regression (SVR)

The notion of the SVR is similar to the Support Vector Machine (SVM) with some minor differences. An SVM classifiers can achieve classification using margin maximization, expressed in eq. 1.
\begin{align*}
\min_{w,b,\xi} & & \frac{1}{2}w^Tw + c \sum_{i=1}^{l} \xi_i \\
y_i(w^T\phi(x^i) + b) & \geq 1 - \xi_i
\end{align*}

where, \(w=\) weight vector, \(b=\) scalar, \(\phi(x^i)=\) optimization problem, \(\xi_i=\) slack variables, \(x^i=\) input vector, \(c=\) related to slack variable.

If the training vector becomes higher and \(w\) is high dimensional then optimization problem in eq. 1 becomes difficult to solve. Hence, Lagrange multipliers [23] are used to solve the optimization problem which are shown in eq. 2.

\[
w = \sum_{i=1}^{l} y^i \alpha_i \phi(x^i)
\]

During the training phase, decision function for any test pattern \(x\) is shown in eq. 3.

\[
f(x) = \text{sgn}(w^T\phi(x) + b) = \text{sgn}(\sum_{i=1}^{l} y^i \alpha_i k(x^i,x) + b)
\]

where, \(\alpha_i=\) Lagrange multipliers also known as Kuhu-Tucker coefficients, \(k(x^i,x) = \phi(x^i)\phi(x)\) is known as the kernel function.

The prediction task is performed by \(\varepsilon\)-insensitive loss function, where the \(\varepsilon\) parameter is used to measure the tolerance of actual \(y\) values. The \(\varepsilon\) tolerates the difference between actual \(y\) and predicted \(\hat{y}\). The optimization problem solved by \(\varepsilon\)-SVR regression is shown in eq. 4.

\[
\begin{align*}
\min_{w,b,\xi} & & \frac{1}{2}w^Tw + c \sum_{i=1}^{l} \xi_i + \sum_{i=1}^{l} \xi_i^* \\
w^T\phi(x^i) + b - y^i & \leq \varepsilon + \xi_i \\
y^i - w^T\phi(x^i) - b & \leq \varepsilon + \xi_i^*
\end{align*}
\]

\[
\xi_i, \xi_i^* \geq 0, i = 1, \ldots, l
\]

Similar to the SVM, Lagrange multipliers also helps in optimization problem solved in SVR and given by eq. 5.

\[
f(x) = \sum_{i=1}^{l} (-\alpha_i + \alpha_i^*) K(x^i,x) + b
\]

There is different kernel function which can be used to set the pattern. The use of appropriate kernel function is essential to set the pattern for non-linear sets. In this proposed approach, sigmoid function has been used for setting the pattern. The different kernel functions with mathematical model are presented in Table 1.

| Kernel function | Mathematical expression |
|-----------------|-------------------------|
| Linear          | \(K(x,y) = x \cdot y\)  |
| Polynomial      | \(K(x,y) = (\gamma(x,y)+c0)d\) |
Radial basis function \( \mathbf{K}(x,y) = \exp(-\gamma |x-y|) \)

Sigmoid \( \mathbf{K}(x,y) = \tanh(\gamma (x,y) + c_0) \)

2.2 Particle Swarm Optimization (PSO)

In order to solve complex mathematical problems, the PSO had inspired the formulation of an optimization algorithm, a debate on a flock’s action is presented. In this case, the concept of which point the entire swarm should land is a complicated problem since it depends on many issues, i.e. the maximisation of the availability of food and the minimisation of the risk of the presence of predators [24]. A swarm of birds flying over a place must find a point to land. The viable solution to the optimization issue is represented in the PSO algorithm by the location of the particle in the optimization space. A particle has a velocity vector that defines the direction and velocity of its motion, so that each particle can search for the best particle at the time to find the best value in the solution space and refer to its motion experience. Each particle's position and velocity are randomly distributed at the initial state in the solution space. Then, depending on the known individual optimal value and the global optimal value, the particle dynamically adjusts the position and speed of the particle.

\[
v_{ij}^{t+1} = wv_{ij}^{t} + c_1 r_1^{t} \left( p_{best_{ij}} - x_{ij}^{t} \right) + c_2 r_2^{t} (g_{best} - x_{ij}^{t}) \tag{6}
\]

\[
x_{ij}^{t+1} = x_{ij}^{t} + v_{ij}^{t+1} \tag{7}
\]

where,

\( v_{ij}^{t+1} = \) velocity of particle

\( x_{ij}^{t+1} = \) location of particle

\( p_{best_{ij}} = \) optimum individual value

\( g_{best} = \) global optimum value

\( c_1, c_2 = \) learning factor

\( r_1, r_2 = \) random number whose value is between 0 and 1

\( i = 1, 2, 3, \ldots \ldots, P \) and \( j = 1, 2, 3, \ldots \ldots, n \)

3. Software based smart city applications

Smart cities have an organisation and infrastructure that is inherently stronger. In a dynamic interplay that simplifies daily life for people living and working in the city, all sectors are involved. The cameras at the bus stops can detect how many people are waiting to board; the sensors on the approaching bus know at any given point in time how many people are riding the bus, and how many vacant seats are there on the bus at present.

A sustainable climate, a reduction of carbon dioxide emissions and the use of renewable energy sources are the core elements of a smart city. By extension, by sustainable and intelligent energy management, the aim of smart cities is to reduce greenhouse gas emissions by about 40 percent in coming years [25]. All these practises, according to the theory, take into account the use of new technology. These new technologies will help in promoting the smart city concept in all cities.
The development of smart cities has a positive impact on the society, overall quality of life, mobility, economy, government and environment. In all, smart cities play a vital role in modifying the society by giving it a creative outline with training and inclusion in various aspect. Overall quality of life is upgraded as it affects the cultural wellbeing, safety and health of the society. It also has a positive impact on environment by taking into account sustainable building and urbanism. The resource management has been improved in smart city due to good planning of different projects linked with it [26]. The benefits for smart cities to government is that they have ease of access to data, maximum services can be provided online and ease of infrastructure planning. The consequence of such a mechanism is citizens" involvement in local authorities decision-making. Residents collaboration will decide their needs. An open platform where public data in a digital format can be reprocessed and analysed is the essence of such a government. Smart cities help economy by taking into account new technology to maximise the workload and improve productivity. These cities will develop a culture for enterprise and innovation, enhanced productivity and in the end leading to better local and global connections. This will lead to improvement in infrastructure and technology, efficient transport, and multimodal mobility. The overall benefits of smart city have been explained in fig. 1.

Figure 1. Overall benefits of smart city

The Smart City model opens up new avenues for novel applications. Data collected from these physical spaces are processed and made available in the form of various services which enable applications to use sensed data or to control the physical world. The range of resources and applications used in smart cities have broad applications as shown in fig. 2. These services include sectors such as transportation (smart road networks, connected cars and public transit), public utilities (smart power, delivery of water and gas), education, health and social care, public safety and protection. Emerging software and services are being applied to the above-mentioned sectors, such as citizens daily lives, emergency management, smart buildings, logistics and smart procurement. Smart grid, smart home, protection, building automation, remote health and wellness monitoring, location-aware applications, mobile payments and other machine-to-machine (M2M) applications are the applications for this portfolio, including implementation for the connected city. To provide these services various mobile applications as well as software are developed to meet the requirement.
The development of these software requires some effort as well as planning. The basic requirement in planning the development of the smart city software has been showed in fig. 3. A company to bid for any of smart city software project, it needs to know the basic effort needed to build the software. This will make work easy for the company to do the bidding and manage profit and maintain the quality over time.

**Figure 2.** Different applications of smart city

**Figure 3.** Effort required to design the software used in smart city application
4. Proposed Model

In this paper, a model has been proposed that will help the developers to predict the efforts required to complete the development of a smart city project. This predicted efforts at the initial level of the project will help the developers to make a proper planning, scheduling of different development activities, effective distribution of resource, controlling of budget, and so on. A prediction model is based on some known parameters those will be used to predict unknown parameter. In this paper, three known parameters, i.e., Line of Code (LOC), type of Arduino board, and manpower, have been considered for prediction of efforts in smart city project using IoT (Internet of Things) hardware. The dataset [27] used in this paper was extracted from three different applications (Street Lightning, Smart Parking, and Waste Management), consisting of 20 projects used in various smart cities.

For calculating the effort of each project used in designing the application, first, the manpower must be known. A dataset has been prepared, which consists of all the three known parameters values. With the help of prepared dataset, the proposed model has been formulated. In this proposed hybrid model, LOC, Arduino Board, Manpower are the variables which has been used to train and test the model. Arduino Board is a categorical variable from the collected data of different project; hence this need to be converted in a numeric variable for making the model. For this conversion, one-hot encoding was performed to prepare the database consisting of all numeric values. The overall architecture of the proposed H-PSOSVR (Hybrid – Particle Swarm Optimization Support Vector Regression) model has been shown in fig.4, which comprises of the entire workflow of developing the model.

**Figure 4. Workflow of H-PSOSVR model**

The proposed model has been designed as follows: the considered parameters in PSOSVR program has been given with 20 different project datasets. The maximum iteration times is 300, the particle vector dimension is 3. The value of range $\xi_i$ is [0,0.5], the kernel parameter $\gamma$ is [0.2,7], the maximum speed of vector is (9 0.2 0.4). The first step in this process is to initialize the parameters of particle swarm which consists of particle number, weighting coefficient, learning factors, particle velocity and position. Then the kernel parameter of SVR is encoded according to the position of particles. The model is than trained with training samples. The input parameters change along with the particles. The fitness value has been calculated. The calculated fitness value has been used to evaluate the model validity. If the desired accuracy is achieved than terminate the condition. Otherwise, proceed the iteration. Update the pbest$_{ij}$,gbest$_{ij}$, particle velocity with eq. 6 and particle position with eq. 7. Validate the SVR model with the obtained optimal parameter with the testing dataset. In the later stage, proposed H-PSOSVR model has been compared with PSO and SVR. The single models selected for comparison have been widely used by researchers as a single model for prediction of...
effort [28]. For comparing the performance of proposed H-PSOSVR model against the performance of PSO and SVR model, few evaluation metrics (MSE, MAE, RMSE, and Accuracy) have been identified [29][30].

5. Result and Discussion
This section describes the experimental results obtained based on the evaluation metrics. The evaluation metrics helps in comparing the proposed model with the other baseline model used. The proposed H-PSOSVR model shows good accuracy, capability of handling outliers and noises in the input dataset in comparison to two other models. The obtained values of different evaluation metrics for all the models have been shown in Fig. 5 to 8.

Figure 5. Variation of MSE from H-PSOSVR Model, PSO and SVR

Figure 6. Variation of MAE from H-PSOSVR Model, PSO and SVR

Figure 7. Variation of RMSE from H-PSOSVR Model, PSO and SVR
Figure 7. Variation of Accuracy from H-PSOSVR Model, PSO and SVR

Figure 8. Variation of RMSE from H-PSOSVR Model, PSO and SVR

The lower MSE, MAE, and RMSE value indicates higher accuracy of prediction. It is inferred that SVR has the lowest MSE in comparison to PSO and proposed H-PSOSVR model. In the case of MAE, PSO has the lowest value in comparison to H-PSOSVR and SVR. The RMSE value for SVR and proposed H-PSOSVR model are better than SVR model. In terms of accuracy of the model in effort prediction, proposed H-PSOSVR model shows 94% whereas for PSO, it is 90% and for SVR it is 88%. Finally, it is concluded from the results that proposed hybrid approach is helpful in accurate effort prediction with very less error rate in comparison to the PSO model and SVR model. This H-PSOSVR model helps the developers to make accurate effort prediction for their project before starting the development processes. It also shows that when the software has been developed with the help of different parameters shows good effort. Ultimately, the accurate effort prediction helps in proper resource allocation in different development activities throughout the development life cycle as well as help to design better software.

6. Conclusion
This paper is focused on providing a methodology that will help in planning the smart city IoT based project. Like the software project, smart city IoT based projects also suffer from the improper planning, shortage of budget, slippage of deadline etc. Due to these reasons, most of the projects fail before the actual deployment. Effort has been found as a crucial and striking factor that affects the reason of project failure. The objective of this study was to formulate a model for predicting effort in early stages of project. In the start, some project factors need to be identified those have higher influence on the effort. Line of code, type of board, and manpower have been identified. With the help of PSO and SVR methodology, a model has been proposed by analysing the relation between the variables. Data of 20 IoT based projects have been used to perform the validation of the model. The result obtained from the proposed model compared with the base line models shows that there is an improvement in accuracy. Estimating effort at the initial phase helps in achieving the work in a proper time and budget. It helps in proper planning and identifying the needs of the application development.

References
[1] Simonofski A, Vallé T, Serral E and Wautelet Y 2019 Investigating context factors in citizen participation strategies: A comparative analysis of Swedish and Belgian smart cities Int. J. Inf. Manage. 102011

[2] Kim T hoon, Ramos C and Mohammed S 2017 Smart City and IoT Futur. Gener. Comput. Syst. 76 159–62

[3] Arasteh H, Hosseinnezhad V, Loia V, Tommasetti A, Troisi O, Shafie-Khah M and Siano P 2016 Iot-based smart cities: A survey EEEIC-International Conference on Environment and Electrical Engineering

[4] Gea T, Lamarca M, Paradells J and Roldán D 2013 Smart cities as an application of Internet of Things: Experiences and lessons learnt in Barcelona Seventh International Conference on Innovative Mobile and Internet Services pp 552–7

[5] Mekki K, Bajic E, Chaxel F and Meyer F 2019 A comparative study of LPWAN technologies for large-scale IoT deployment ICT Express 5 1–7

[6] Altulyan M, Yao L, Kanhere S S, Wang X and Huang C 2019 A unified framework for data integrity protection in people-centric smart cities Multimed. Tools Appl.

[7] AL Zamil M G H, Samarah S, Rawashdeh M, Karime A and Hossain M S 2019 Multimedia-oriented action recognition in Smart City-based IoT using multilayer perceptron Multimed. Tools Appl. 78 30315–29

[8] Iqbal M M, Farhan M, Jabbar S, Saleem Y and Khalid S 2019 Multimedia based IoT-centric smart framework for eLearning paradigm Multimed. Tools Appl. 78 3087–106

[9] Ryan BJ Everimpact. https://www.everimpact.org/#home

[10] Telensa. https://www.telensa.com/

[11] Luminext. https://www.luminext.eu/en/

[12] Echelon. https://www.echelon.com/applications/pl-rf-outdoor-lighting

[13] Bestmile. https://bestmile.com/

[14] Kiunsys. https://www.kiunsys.com/

[15] Anagog. https://www.anagog.com/

[16] Streetline. https://www.streetline.com/parking-analytics/

[17] Enevo. https://www.enevo.com/

[18] Bigbelly. http://bigbelly.com/

[19] Compology. https://compology.com/

[20] Ayyeka. https://www.ayyeka.com/wavelet-kits.html#wastewater

[21] Schloss B, Srivastava B and Uceda-Sosa A R 2011 SCRIBE Models and Methodology: Smart Cities Reference Information and Behavior Exchange ontologies

[22] Uceda-Sosa R, Srivastava B and Schloss R J 2011 Building a highly consumable semantic model for smarter cities Proceedings of the AI for an Intelligent Planet on - AIIP pp 1–8

[23] Rockafellar R . T 1993 Lagrange Multipliers and Optimality SIAM Rev. 35 183–238

[24] Fernandes Junior F E and Yen G G 2019 Particle swarm optimization of deep neural networks architectures for image classification Swarm Evol. Comput. 49 62–74

[25] Meijer A and Bolívar M P R 2016 Governing the smart city: a review of the literature on smart urban governance Int. Rev. Adm. Sci. 82 392–408

[26] Kim T hoon, Ramos C and Mohammed S 2017 Smart City and IoT Futur. Gener. Comput. Syst. 76 159–62

[27] (2018) ISBSG
[28] Nassif A B, Azzeh M, Capretz L F and Ho D 2016 Neural network models for software development effort estimation: a comparative study Neural Comput. Appl. 27 2369–81

[29] Jorgensen M and Moløkken-Østvold K 2004 Reasons for software effort estimation error: Impact of respondent role, information collection approach, and data analysis method IEEE Trans. Softw. Eng. 30 993–1007

[30] Murillo-Morera J, Quesada-López C, Castro-Herrera C and Jenkins M 2016 An Empirical Validation of an Automated Genetic Software Effort Prediction Framework Using the ISBSG Dataset CIBSE 2016 - XIX Ibero-American Conf. Softw. Eng. Quito, Ecuador 1–15