Theoretical and Probabilistic Conversion Model between FPA and COSMIC Measurement Method

Abedallah Zaid Abualkishik (azasoft1@gmail.com)
Mohd Hasan Selamat (mhbselamat@gmail.com)
Abdul Azim Abd Ghani (azim@fsktm.upm.edu.my)
Rodziah Atan (rodziah@fsktm.upm.edu.my)
Department of Information System, Faculty of Computer Science and Information Technology, University Putra Malaysia, 43400 UPM Serdang, Selangor Darul Ehsan, Malaysia

Jean-Marc Desharnais (desharnaisjm@gmail.com)
École de Technologie Supérieure (ÉTS) at the Software Engineering and IT Department, 1100 Notre-Dame West, Montréal, Canada.

Adel Khelifi (a.khelifi@alhosnu.ae)
Software Engineering Department,
ALHOSN University, P.O. Box 38772, Abu Dhabi, UAE

Introduction

- After the COSMIC method has been proposed, many convertibility studies [1], [2] ,[3], [4], [5], [6], [7], [8], [9] have been published to tackle the issue of converting the size measured by FPA method, i.e. UFP to its corresponding CFP size.

- In the literature, its possible to classify the conversion types between FPA and COSMIC into Statistical and Theoretical conversion. The first converts the FPA data to the corresponding COSMIC data statistically using different regression models. The second converts the FPA data to COSMIC data based on conceptual correspondence, i.e. rules, concepts and schematics of the two methods.

Introduction

- In the literature, there is only one theoretical conversion model proposed in [13].

- The model uses the number of File Type Referenced (FTR) accessed in each FPA elementary process to predict the corresponding number of data movements in the corresponding functional process. However, the model might produce invalid results when applied to another dataset due to some shortcomings in the proposed conversion equations.

- The model produce a COSMIC interval size where the actual size will lie within the following interval:

$$\sum_{i=1}^{EI} MAX(2, FTRi + 1) + \sum_{i=1}^{EO} MAX(2, FTRi + 1) + \sum_{i=1}^{EQ} MAX(2, FTR + 1)$$
$$\leq CFSU \leq$$
$$\sum_{i=1}^{EI} MAX(2, 2 * FTRi + 1) + \sum_{i=1}^{EO} MAX(2, 2 * FTRi + 1) + \sum_{i=1}^{EQ} MAX(2, 2 * FTRi + 1)$$

- However, the proposed model is subject to some uncertainties (mapping issues) and might produce invalid results, for example, Fetcke case study.

- There are many unconsidered cases, flaws and limitations in the proposed model, such as:
- The model does not take into account the Entry data movement of the COSMIC method in the External Output and External Inquiry elementary processes of the FPA method.
- The model does not take into account the Write data movement of the COSMIC method in the External Output elementary process of the FPA method.
- The model does not take into account the differences between the ILF and EIF referenced in the elementary process to predict the COSMIC data movements.
- The model assigns an equal value for the minimum and maximum bound to all TF types.
- The number of COSMIC data movements assumed for the minimum bound of the External Inquiry violates the processing logic of IFPUG FPA definition of the External Inquiry.
- The model assigned an error message for the minimum bound for each functional process. However, it might not be the case for all the functional processes.

- A concrete example (**Fetcke Case Study**)

No.	Elementary Process	FTR	FP	Actual CFP	Min	Max
1	Add customer	1	3	4	2	3
2	Change Customer data	2	3	8	3	5
3	Delete customer	2	3	5	3	5
4	Receive payment	1	3	4	2	3
5	Deposit item	3	6	6	4	7
6	Retrieve item	2	3	6	3	5
7	Add place	1	3	4	2	3
8	Change place data	2	3	8	3	5
9	Delete place	2	3	5	3	5
10	Print customer item list	2	5	5	3	5
11	Print bill	2	4	5	3	5
12	Print stored item list	1	4	3	2	3
13	Query customers	1	3	4	2	3
14	Query customers items	2	4	5	3	5
15	Query places	1	3	4	2	3
16	Query stored items	2	3	5	3	5
	Total	27	77	81	43	70

A New Theoretical Proposed Conversion Model

- **FPA TFs Analysis According to COSMIC Viewpoints:**

For each ILF referenced:

- External Input (EI): Def requires (E, W). However, it could have the four COSMIC data movements.
- External Output (EO): Def requires (E, R, X). However, it could have the four COSMIC data movements.
- External Inquiry (EQ): Def “strictly” requires (E, R, X) only.

- By definition, an EIF is a single file that locates outside application boundary. The Def “strictly” requires (E, X) only over the persistent storage.

Minimum and Maximum Equations:

The IFPUG-CPM [18] determined the primary intent for each transaction type as shown in the following table:

Function:	TF BFC type :		
	EI	EO	EQ
Alter the behavior of the system	PI	F	N/A
Maintain one or more of ILFs	PI	F	N/A
Presents information to a user	F	PI	PI

Transaction type accesses single ILF	E	R	W	X
EI	1 (Trigger or data attribute)	0	1	0
EO and EQ	1 (Trigger)	1	0	1
Transaction type that does not access any file.				
EO	1	0	0	1

$$\text{Minimum CFP} = \sum_{i=1}^{M} \sum_{j=1}^{N} \text{EI}i((\text{ILF}j + 1) + 2\text{EIF}j)) + \sum_{i=0}^{M} \sum_{i=1}^{N} \text{EO}i((\text{ILF}j + 2) + 2\text{EIF}j))$$

$$+\sum_{i=0}^{M} \sum_{j=1}^{N} \text{EQ}i((2\text{ILF}j + 1) + 2\text{EIF}j)) \quad (1)$$

$$\text{Maximum CFP} = \sum_{i=1}^{M} \sum_{j=1}^{N} \text{EI}i((4\text{ILF}j + 1) + 8\text{EIF}j)) + \sum_{i=0}^{M} \sum_{i=1}^{N} \text{EO}i((4\text{ILF}j + 1) +$$

$$8\text{EIF}j)) + \sum_{i=0}^{M} \sum_{j=1}^{N} \text{EQ}i((3\text{ILF}j + 1) + 8\text{EIF}j)) \quad (2)$$

IFPUG TF BFCs processing logic forms

Form of processing logic	TF type:		
	EI	EO	EQ
Validation are performed	c= DM	c=DM	c=DM
Mathematical formula and calculations are performed	c=R, DM	m*=R, DM	n
Equivalent values are converted	c=R	c=R	c=R
Data is filtered and selected by using specific criteria to compare multiple sets of data.	c=R, DM	c=R, DM	c=R, DM
At least one ILF is updated	m*=W	m*=W	n
At least one ILF or EIF is referenced	c=ILF(R), EIF (E,X)	c=ILF(R), EIF (E,X)	m=ILF(R), EIF (E,X)
Data or control information is retrieved	c=R	c=R	m=R
Derived data is created	c=R, DM,X	m*=R,DM,X	n
Behavior of the system is altered	m*=W,X	m*=W,X	n
Prepare and presents information outside the boundary	c=R, X	m=R,X	m=R,X
Capability to accept data or control information that enters the application boundary.	m=E	c=E	c=E

Probabilistic Minp and Maxp Equations

- Basically, the Minp equation considers
  (the data movements of equation (1), i.e. Primary Intent (PI)) +
  (the probable number of occurrences for certain data movements) +
  (the fixed Entry according to the definition of COSMIC functional process.)

- The Maxp equation assumes full coverage of equation (1) +
  Other forms of processing logic (probability theory principles) +
  Error message (X).

- Expected value assumes uniform distribution. E.g, N=4. The probability of having (1 or 2 or 3 or 4) are emerged equally.

- Statistically, the total weighted probabilities value of one variable is equal to $\left(\frac{N}{2} + 0.5\right)$

- Minp equation assumes 1 time (2 data movements) for EIF. Maxp 3 times (6 data movements)

Probable application size

The proposed equation estimate a probable interval of minimum and maximum bounds by means of the following formula:

$$\sum_{i=1}^{M}\sum_{j=0}^{N} EIi\left(\left(1.5+\frac{3ILFj}{2}\right)+(2EIFj)\right)+\sum_{i=1}^{M}\sum_{j=0}^{N} EOi\left(\left(2.5+\frac{3ILFj}{2}\right)+(2EIFj)\right)+$$
$$\sum_{i=1}^{M}\sum_{j=0}^{N} EQi((1+2ILFj)+(2EIFj))$$

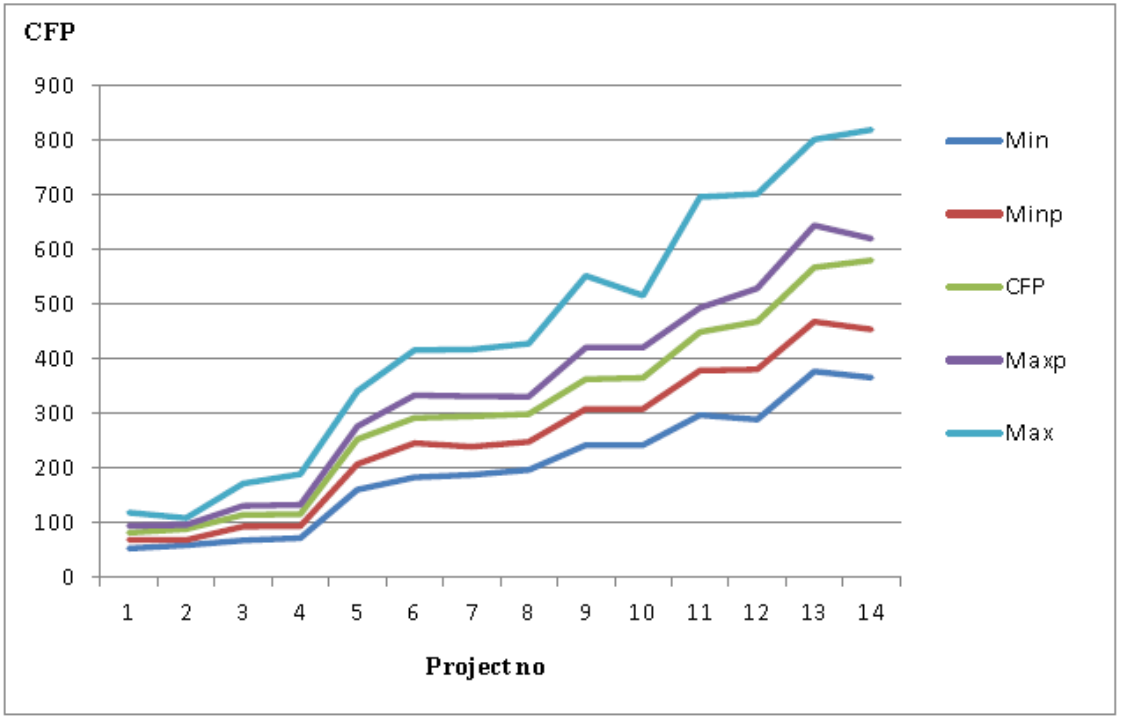
$$\leq \mathbf{CFP} \leq$$

$$\sum_{i=1}^{M}\sum_{j=0}^{N} EIi\left(\left(1.5+\frac{5ILFj}{2}\right)+(6EIFj)\right)+\sum_{i=1}^{M}\sum_{j=0}^{N} EOi\left(\left(2.5+\frac{5ILFj}{2}\right)+(6EIFj)\right)+$$
$$\sum_{i=1}^{M}\sum_{j=0}^{N} EQi((2+2ILFj)+(6EIFj)) \qquad (3)$$

Direct validation

- The dataset published in [5] used to validate the proposed equation.

ID	Minimum	Minp	CFP	Maxp	Maximum
1	52	68.5	81	93.5	118
2	58	67.5	88	95.5	108
3	67	92.5	114	130.5	171
4	71	93	115	132	188
5	160	206.5	252	275.5	340
6	182	245	291	332	415
7	187	238.5	294	330.5	416
8	196	247	298	330	427
9	241	307.5	362	419.5	551
10	241	307.5	364	419.5	515
11	296	377.5	448	492.5	695
12	288	379.5	467	527.5	700
13	376	467	566	643	800
14	365	453	579	619	818



Statistical validation

- D1= CFP-Minp.
- D2= Maxp-CFP

ID	D1	D2
1	12.5	12.5
2	20.5	7.5
3	21.5	16.5
4	22	17
5	45.5	23.5
6	46	41
7	55.5	36.5
8	51	32
9	54.5	57.5
10	56.5	55.5
11	70.5	44.5
12	87.5	60.5
13	99	77
14	126	40

The null hypothesis of this test, *Ho*, is that the defined random variables (D1 and D2) follow an **exponential** distribution.

Kolmogorov-Smirnov test results

	D1	**D2**
Most Extreme Absolute Differences	**0.27**	**0.21**
Asymptotically Significant "p-value"	**0.23**	**0.51**
α	**0.01**	**0.01**

Conclusion

- The statistical conversion does not consider the differences between FPA BFCs and their corresponding CFP results.
- In this paper, a set of conversion equations have been proposed based on the primary intents and processing logic forms of FPA transaction function BFCs, COSMIC method rules and measurement good practices.
- The model results prove the effectiveness of the transaction functions conversion.

- Given the FPA measurement details at the level of transaction function BFCs and it is associated files accessed, the proposed equations estimate the minimum and maximum bounds where the CFP size will fall into. The Minp and Maxp equations estimate an interval where the CFP is most likely will fall into.
- The model has been validated directly and statistically using 14 industrial business applications. The model produced high accurate results for the tested cases.

Q & A