REVUE DE STATISTIQUE APPLIQUÉE

J. L. BON
M. BOUISSOU

Fiabilité des grands systèmes séquentiels : résultats théoriques et applications dans le cadre du logiciel GSI

Revue de statistique appliquée, tome 40, n° 2 (1992), p. 45-54
<http://www.numdam.org/item?id=RSA_1992__40_2_45_0>

Rev. Statistique Appliquée, 1992, **XXXX** (2), 45-54

FIABILITÉ DES GRANDS SYSTÈMES SÉQUENTIELS : RÉSULTATS THÉORIQUES ET APPLICATIONS DANS LE CADRE DU LOGICIEL GSI

J.L. BON

Laboratoire de Probabilité Statistique,
Université Paris-Sud, bâtiment 425, 91405 Orsay cedex

M. BOUISSOU

Département ESF, EDF-DER, 1 av. du Général de Gaulle,
92141 Clamart cedex, Tél : (1) 47 65 58 22

Introduction

Les études de fiabilité des systèmes s'appuient sur deux grandes catégories de méthodes : la simulation de Monte Carlo, et les méthodes analytiques. La deuxième catégorie se ramifie très finement en fonction des types de systèmes étudiés, cependant on peut y distinguer deux grandes familles, qui sont :

- les méthodes booléennes, ou combinatoires,
- les méthodes markoviennes.

Les méthodes combinatoires sont employées pour les systèmes dont le comportement n'est pas séquentiel, ce qui revient à dire qu'il est raisonnable de supposer leurs composants indépendants les uns des autres, alors que les méthodes markoviennes sont préférées pour les systèmes séquentiels, à condition toutefois que l'on puisse raisonnablement supposer qu'ils «oublient» tout le passé à chaque instant.

Si les méthodes combinatoires sont à même aujourd'hui de traiter des systèmes de grande taille (plusieurs centaines de composants), les méthodes markoviennes, elles, sont généralement limitées à une quinzaine de composants car elles font appel à la construction du graphe des états du système, dont la taille croît de manière exponentielle en fonction de la taille du système. Cette limite en nombre de composants ne peut être franchie que par l'exploitation de spécificités du système, qui permettent d'agréger des états.

L'objet de cet article est d'apporter une contribution à l'étude des systèmes séquentiels généraux, qui satisfont une hypothèse un peu moins forte que celle de Markov. Dans une première partie, nous allons donner un encadrement effectivement calculable sur des sytèmes de grande taille, de la fiabilité du système : cet encadrement est d'autant plus précis que le système est fortement et rapidement réparable (cette notion sera précisée ci-après).

Nous montrerons dans une deuxième partie de l'article comment ces résultats théoriques peuvent donner lieu à de nombreuses applications pratiques grâce au logiciel GSI, un des éléments de l'atelier d'analyse de la sûreté de fonctionnement FIGARO, développé à EDF [5]. Cet atelier permet par ailleurs l'utilisation conviviale de toutes les méthodes évoquées ci-dessus, grâce à ses possibilités de génération automatique des modèles fiabilistes à partir de saisies graphiques d'information.

Intérêt d'une approximation exponentielle de la fiabilité

Avant d'établir les résultats obtenus, nous résumons les travaux antérieurs mettant en relief le rôle particulier joué par la distribution exponentielle du temps avant défaillance d'un système «fortement réparable». On pourra consulter l'article de J. Keilson [1] pour les démonstrations.

Nous simplifions la description des états du système physique en ne considérant que les trois ensembles d'états suivants :

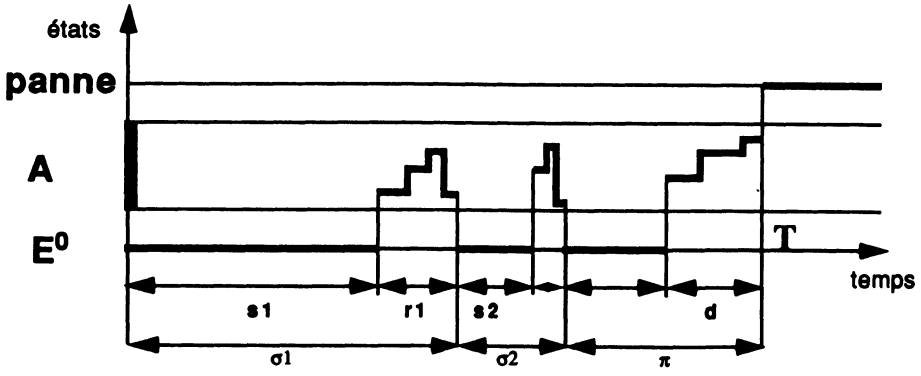
– l'état initial de fonctionnement parfait : E^0,

– l'ensemble des états de fonctionnement différents de E^0 : A,

– l'état E^∞, absorbant, qui regroupe tous les états de panne du système.

Le fonctionnement du système est alors décrit par le processus de renouvellement associé aux dates successives t_i auxquelles le système retrouve l'état parfait.

Supposons que lorsque le système entre dans A, il a la probabilité α de parvenir à l'état de panne et $1 - \alpha$ de revenir à l'état initial E^0.



Notons ν le nombre de sorties de l'état initial E^0. La probabilité d'avoir exactement k sorties est :

$$P(\nu = k) = \alpha.\ (1 - \alpha)^{k-1}$$

Soit T La date d'apparition de la panne ; soit π le temps séparant la dernière arrivée en E° de la panne. Alors, pour $\nu = k$ nous pouvons décomposer T de la façon suivante :

$$T = \sigma_1 + \sigma_2 + \dots \sigma_{k-1} + \pi$$

Toutes les variables σ_i ont la même loi qu'une variable σ représentant la durée entre deux arrivées successives en E^0 ; les variables σ_i et π sont toutes indépendantes.

Théorème 1 : Désignons par $E(T)$ le temps moyen jusqu'à la panne (ce que l'on appelle MTTF), $E(\sigma)$ le temps moyen de retour, $E(\pi)$ le temps moyen de panne sans retour à l'état initial, alors :

$$E(T) = E(\pi) + \frac{(1-\alpha)}{a} \cdot E(\sigma)$$

ce que l'on peut interpréter de la façon suivante : le temps moyen de panne est la somme du temps moyen de panne sans retour et de la durée moyenne cumulée des retours.

La démonstration de ce théorème résulte de l'étude de la transformée de Laplace du temps T de panne qui s'écrit par conditionnement sur ν :

$$L(T) = \sum_{k=1}^{k=\infty} L(\sigma)^{k-1} \cdot L(\pi) \cdot \alpha \cdot (1-\alpha)^{k-1}$$

(le symbole L désigne la transformée de Laplace)

D'où l'on tire :

$$L(T) = \alpha \cdot L(\pi)/(1 - \beta \cdot L(\sigma)) \qquad \text{avec } \beta = 1 - \alpha$$

Le calcul du temps moyen est obtenu par dérivation en O.

Le théorème ci-dessous justifie l'utilisation fréquente de la distribution exponentielle pour la durée de vie des systèmes «bien réparables». Nous supposons en effet que le temps moyen de panne sans retour $E(\pi)$ est négligeable devant la durée moyenne cumulée des retours $(\beta/\alpha) \cdot E(\sigma)$

Théorème 2 : Si α tend vers 0 avec $E(\pi) << (\beta/\alpha) \cdot E(\sigma)$ alors :

$$P(T \geqslant x) \text{ tend vers } \exp\left(\frac{-x}{E(T)}\right)$$

C'est en utilisant la décomposition de la transformée de Laplace ci-dessus que Keilson obtient cette approximation exponentielle.

Ce résultat permet des applications numériques rapides pour les systèmes «fortement réparables» (c'est-à-dire pour lesquels la probabilité α est faible), mais ne contient aucune précision sur l'erreur faite. La fiabilité ainsi calculée peut être exagérement optimiste.

Nous avons cherché à exhiber une approximation exponentielle proche de la fiabilité réelle mais fournissant une valeur pessimiste.

Approximation exponentielle «sans retour à l'état initial» (SRI)

L'observation d'un système «fortement réparable» amène à constater que l'essentiel du temps de fonctionnement est passé dans l'état parfait E^0. Dans le modèle précédent nous pouvons chercher à mettre en évidence les différents temps de séjour dans E^0.

Soit s la durée d'un séjour dans E^0, nous pouvons décomposer les variables σ et π sous la forme :

$$\sigma = s + r$$
$$\pi = s + d$$

La variable r désigne le temps mis à revenir à l'état inital E^0 après l'avoir quitté, conditionnellement au fait que l'état de panne n'est pas atteint. La variable d désigne le temps mis à atteindre la panne à partir de la sortie de E^0 sachant l'état de panne atteint.

Les formules précédentes s'écrivent maintenant pour $\nu = k$:

$$T = s_1 + s_2 + \ldots s_k + r_1 + r_2 + \ldots + r_{k-1} + d$$

et pour les espérances, nous avons la relation suivante :

$$E(T) = \frac{1}{\alpha}E(s) + \frac{1-\alpha}{\alpha}E(r) + E(d)$$

Pour les systèmes «rapidement réparables», le temps de bon fonctionnement est grand par rapport au temps de réparation ou de panne et nous sommes amenés à étudier l'approximation naturelle : $T \approx s_1 + s_2 + \ldots s_\nu, \quad E(T) \approx \frac{1}{\alpha}E(s)$

Un premier avantage à cette approximation est qu'elle est pessimiste : le système est considéré comme moins fiable qu'il ne l'est réellement. De plus, il nous a été possible d'obtenir un majorant de l'erreur commise.

Désignons par S la somme des s_i représentant les temps de séjour successifs dans E^0. Le théorème ci-dessous précise la qualité de l'approximation obtenue :

Théorème 3 : Soit ν le nombre de périodes passées dans l'état E^0, soient s, r, d les différents temps définis ci-dessus ; T désigne le temps de panne. Soit X le processus de renouvellement ne prenant en compte que les séjours dans E^0 : $X_n = s_1 + s_2 + \ldots s_n$

Soit H la fonction de renouvellement du processus X pour lequel les interarrivées s ont chacune la répartition F :

$$H(t) = \sum_{n=1}^{\infty} F^n(t)$$

où F^n est la convolution de n fonctions F (au sens de Laplace-Stieljes)

Si H est sous-additive, alors la fiabilité du système $P(T \geqslant x)$ est encadrée de la façon suivante :

$$P(S \geqslant x) \leqslant P(T \geqslant x) \leqslant P(S \geqslant x) + \alpha \cdot E(H(d)) + (1 - \alpha) \cdot E(H(r))$$

La première inégalité est évidente et traduit le caractère pessimiste de l'approximation. La seconde s'obtient en conditionnant sur le nombre de sorties et en utilisant les propriétés de sous-additivité de la fonction de renouvellement. Cette formule s'applique en particulier pour des fonctions de répartition F correspondant à des taux de sortie de E^0 décroissants. Un cas intéressant est obtenu pour un temps de séjour s de loi exponentielle. On retrouve alors l'une des inégalités de Soloviev données dans [2] :

Corollaire : Si chaque temps de séjour dans E^0 est exponentiel de paramètre Λ, la formule ci-dessus devient :

$$P(S \geqslant x) \leqslant P(T \geqslant x) \leqslant P(S \geqslant x) + \Lambda \cdot \alpha \cdot E(d) + \Lambda \cdot (1 - \alpha) \cdot E(r)$$

On peut approcher la fiabilité par : $P(S \geqslant x) = \exp(-\Lambda \cdot \alpha \cdot x)$

En effet, dans le cas où la loi F de s est exponentielle, la fonction de renouvellement s'écrit : $H(u) = \Lambda \cdot u$ et la somme des ν variables exponentielles (où la loi de ν est géométrique) est elle-même exponentielle de paramètre $\Lambda \cdot \alpha$.

On constate que cette approximation est d'autant plus précise que le système est plus *fortement* (α petit devant 1) et *rapidement* ($E(r)$ et $E(d)$ petits devant $1/\Lambda$) *réparable*. On peut aussi remarquer que la majoration n'a guère d'intérêt pour les faibles valeurs de x. Dans ce cas, il est préférable d'utiliser un autre encadrement dû à Soloviev [2] :

Théorème 4 : Avec les hypothèses précédentes, en particulier F exponentielle de paramètre Λ, la fiabilité du système vérifie :

$$\exp(-\Lambda\cdot\alpha\cdot x) \leqslant P(T \geqslant x) \leqslant \exp(-\Lambda\cdot\alpha\cdot x)+\Lambda\cdot\alpha\cdot E(d)+\Lambda^2\cdot\alpha\cdot(1-\alpha)\cdot x\, E(r)$$

pour les «petites» valeurs de x.

Utilisation dans le cas Markovien

Nous supposons dorénavant que le fonctionnement du système physique est bien décrit par un processus markovien. L'état initial est E^0. Le temps de séjour dans E^0 suit la loi exponentielle de paramètre Λ égal à la somme des taux de transition sur toutes les sorties possibles i de E^0; Ces états i sont appelés les amorces de panne.

Le modèle précédent s'adapte bien à ce cas. Nous obtenons l'encadrement suivant :

$$\exp\left(-\sum_i \lambda_i \cdot \alpha_i \cdot x\right) \leqslant P(T \geqslant x) \leqslant \exp\left(-\sum_i \lambda_i \cdot \alpha_i \cdot x\right) + \sum_i \lambda_i(\alpha \cdot E(d_i) + (1-\alpha)E(r_i)), \text{ où :}$$

λ_i est le taux de transition de E^0 vers l'amorce i,

α_i est la probabilité d'atteindre la panne avant E^0 en partant de l'amorce i,

r_i est le temps fini de retour en E^0 à partir de l'entrée en i,

d_i est le temps fini de panne à partir de l'entrée en i.

On peut affiner le modèle précédent en prenant en compte le temps passé dans les amorces, puisqu'il s'agit de points de renouvellement. Les calculs analytiques deviennent alors très techniques et nous donnons le résultat obtenu :

Théorème 5 : (approximation sans retour et avec amorces : SRIA)

La fiabilité à l'instant x est minorée par : $\exp\left(-\sum_i \lambda_i \cdot \alpha_i \cdot g_i(x)\right)$, avec :

$$g_i(x) = (\exp(-A_i \cdot x) - 1 + A_i \cdot x)/A_i$$

A_i = somme des taux de sortie de l'amorce i.

Ces formules ont été appliquées sur des petits systèmes (quelques composants) afin de confirmer leur intérêt pratique (précision des approximations). Nous donnons ci-après un exemple illustrant leur caractère opérationnel, et le gain qu'elles apportent par rapport à des méthodes déjà connues *sur un cas d'étude réel.*

Applications pratiques dans le cadre du logiciel GSI

Pour la classe des systèmes markoviens, les calculs analytiques peuvent reposer soit sur des méthodes matricielles, (mais cela limite le nombre de composants du système à une quinzaine) soit sur des méthodes séquentielles.

L'article [4] donne une méthode par séquences qui permet le calcul d'un encadrement de la fiabilité, précis et applicable en pratique, pour des systèmes beaucoup plus importants, à condition *qu'ils soient non réparables, ou qu'on s'intéresse à des temps de mission suffisamment «petits»*. Les approximations plus globales telles que celles présentées ci-dessus permettent d'aller bien plus loin dans des situations complémentaires.

Le logiciel GSI, développé à EDF à partir de 1985 [4], intègre plusieurs types de calculs, afin de profiter de leur complémentarité, et de permettre différents compromis entre précision souhaitée et finesse de modélisation.

A partir d'une description unique du système sous forme de règles de production, il permet de calculer la fiabilité, grâce à des calculs matriciels, ou à une exploration systématique des séquences qui peuvent amener le système de l'état initial à un état de panne, *sans construire le graphe.*

Cette méthode permet de transformer le problème rédhibitoire de l'explosion du nombre d'états en un problème beaucoup plus maîtrisable d'explosion du nombre de séquences à explorer.

En effet, il s'avère qu'en pratique, sur les systèmes réels, on peut obtenir des estimations précises de la fiabilité et de la disponibilité asymptotique du système en n'explorant qu'un nombre limité de séquences, en fonction d'heuristiques, telles que le choix des séquences d'une probabilité minimale, et d'une longueur maximale (par «longueur», nous entendons «nombre de transitions»).

GSI a d'abord servi d'outil de recherche pour explorer diverses voies d'étude des systèmes séquentiels complexes [6], puis a été repris sous forme d'un logiciel développé sous assurance qualité, et utilisé par EDF et le CEA pour de nombreuses évaluations dans le cadre des études probabilistes de risque des centrales nucléaires 900MW et 1300MW d'EDF.

Le langage d'entrée de GSI, sous forme de règles de production auxquelles on associe une sémantique temporelle, est déjà de haut niveau par rapport à la description «développée» d'un graphe. En effet, quelques lignes de langage GSI peuvent résumer un graphe énorme, voire infini.

Cependant, la description en règles d'un système comportant plus d'une dizaine de composants peut s'avérer fastidieuse et génératrice d'erreurs. Il est en particulier dommage de ne pouvoir réutiliser les règles que l'on a écrites pour un système dans une autre étude.

Ainsi s'est fait sentir la nécessité d'un langage de modélisation de plus haut niveau que le langage GSI, qui garderait la même généralité et la même souplesse, mais qui permettrait en outre de décrire des connaissances génériques sur des classes de systèmes.

C'est suite à ce constat qu'a été développé le **langage FIGARO** de modélisation des systèmes, qui présente les deux importantes caractéristiques suivantes : il constitue une *généralisation naturelle* de la plupart des modèles fiabilistes classiques, et il permet la constitution de *bases de connaissances génériques* sur des classes de systèmes.

Le langage FIGARO peut être exploité par l'utilisateur au sein d'un atelier aux multiples possibilités, très convivial grâce à des interfaces graphiques évoluées. Cet atelier permet notamment la génération automatique de modèles spécifiques d'un système (arbre de défaillance, modèle pour GSI, réseau de Petri stochastique), à partir de connaissances génériques sur le système, décrites en langage FIGARO, et du schéma du système, saisi graphiquement.

L'atelier FIGARO et ses principaux concepts sont décrits dans l'article [5] ; nous allons maintenant revenir aux applications pratiques des résultats donnés au début de cet article, tout en les replaçant par rapport à d'autres méthodes.

Nous allons illustrer ci-après les différents types de calculs autorisés par GSI, à partir d'un exemple **tiré d'une étude réelle** de système complexe. Cet exemple peut être considéré comme très représentatif des modèles de systèmes réparables redondants que l'on peut avoir à évaluer, à la fois sur le plan de la structure du graphe (qui tient compte d'une redondance d'ordre 4, avec des possibilités de défaillances de cause commune sur des paires de composants) et des données de fiabilité ; les transitions du graphe sont soit des défaillances, de taux compris entre 9e-10 et 2.9e-3 h^{-1}, soit des réparations, de taux compris entre 1.7e-3 et 0.37 h^{-1}.

Cet exemple, bien que réaliste, est très petit (14 états, 53 transitions), ce qui nous a permis de pousser les calculs par séquences à un degré de précision maximum, et de les comparer avec le calcul matriciel, qui nous sert ici de référence.

Le tableau en annexe donne diverses indications sur la précision et le coût, mesuré par le temps de réponse sur une station de travail SUN (sparcstation 1), des trois principaux modes de calcul que GSI permet.

Dans le calcul dit «par séquences normales» (lignes b et c), GSI explore et quantifie *toutes* les séquences (ou trajectoires du processus) qui amènent le système de l'état initial à l'état de panne, tout en ayant une probabilité minimale de réalisation *sur le temps de mission considéré*, et une longueur maximale. La somme des probabilités des séquences incomplètement explorées fournit un majorant de ce qui est négligé. Ce calcul est décrit dans l'article [4].

Par un principe analogue, mais en explorant cette fois les séquences allant de l'état initial à la panne sans repasser par l'état initial, *dans la chaîne de Markov immergée* du graphe étudié, GSI permet le calcul d'un encadrement de la valeur α de la partie théorique du présent article.

Ici, il a été possible d'obtenir α avec une précision relative meilleure que 1e-4, en 9 secondes de calcul seulement.

La valeur du minorant $b1$ de la défiabilité n'est qu'une approximation, obtenue à partir de la formule du théorème 4 (qui est plus précise que celle du théorème 3 pour les temps de mission considérés), en estimant les grandeurs $E(d)$ et $E(r)$ par la somme des durées des séquences amenant directement à la panne ou retournant à l'état initial, pondérées par leurs probabilités.

Le tableau en annexe illustre bien le fait que lorsque l'on ne peut faire appel au calcul matriciel à cause du gigantisme de l'espace des états, ou parce que ce calcul est trop long (c'est le cas, même pour un petit graphe, si l'on s'intéresse à un temps de mission trop grand), on peut adopter la démarche suivante :

lancer un calcul par séquences normales, qui donne un premier encadrement de la défiabilité; si cet encadrement est trop large (c'est dû au fait que l'on s'intéresse à un temps de mission trop élevé pour un système dont le graphe comporte des cycles), essayer le calcul SRI, qui permettra sans doute d'affiner la borne supérieure de la défiabilité. Si l'encadrement reste peu satisfaisant, la seule solution consiste à essayer d'estimer l'erreur de l'approximation SRI au mieux par les formules des théorèmes 3 et 4.

Par ailleurs, l'on voit que l'approximation SRI est la seule estimation raisonnablement accessible au delà d'un certain temps de mission pour un système «très réparable».

Références

[1] KEILSON J., «Stochastic models in reliability theory» in Proc. of Int. School of Physic 1986 North-Holland

[2] SOLOVIEV A.D., «Méthodes analytiques...», in GNEDENKO B;V Voprosi matematitcheskoi teorii nadejnosti 1983 Radio Viaz (en russe).

[3] BON J.L., BRETAGNOLLE J., CHAUVEAU D., JAKUBOWICZ P., PAMPHILE P., RAOULT J.P., «Calcul par séquences sur les graphes d'états» : rapport du C.E.M.S. Equipe de statistique appliquée d'Orsay. Université Paris-Sud. Mathématiques (Bâtiment 425) 91405 ORSAY.

[4] M. BOUISSOU, «Recherche et quantification automatiques de séquences accidentelles pour un système réparable», 5ème Congrès de Fiabilité et Maintenabilité de Biarritz, Octobre 1986.

[5] M. BOUISSOU, H. BOUHADANA, M. BANNELIER, N. VILLATTE «Modélisation des connaissances et traitements fiabilistes : présentation du langage FIGARO et des outils associés» Communication au 3ème colloque FIABEX, déc 1990 et note EDF HT-53/90-74A.

[6] N. VILLATTE, «Problèmes posés par les études de fiabilité de systèmes électriques, et automatisation de ces études par des techniques de l'intelligence artificielle», Thèse de doctorat, EDF/Paris VI, Janvier 1990.

ANNEXE

tableau de résultats

Note 1 : dans chaque colonne, les deux valeurs en italiques sont celles qui fournissent le meilleur encadrement de la défiabilité (à condition de faire abstraction du petit problème signalé dans la note 2, ci-après).

Note 2 : le «minorant» calculé par séquences normales (ligne b) est supérieur à la vraie valeur de la défiabilité pour un temps de mission de 100 heures; cette légère anomalie est due à une approximation pessimiste utilisée par GSI pour le calcul de certaines séquences, les formules données dans [4] n'étant pas toujours applicables.

temps de mission (heures)$\Longrightarrow$	10	10^2	10^3	10^4	10^5	10^6
a : calcul matriciel de la *défiabilité*	2.299e-8	2.969e-7	3.064e-6	3.066e-5	3.066e-4	3.062e-3
temps de réponse	7 s	9 s	16 s	45 s	305 s	2940 s
b : minorant par séq. normales	*2.290e-8*	*2.981e-7*	temps de calcul rédhibitoire pour un bon encadrement :			
c : majorant par séq. normales	*2.486e-8*	3.777e-7	le système étant très réparable, au delà d'un certain temps, les probabilités des			
temps de réponse	17 s	476 s	séquences bouclées qui repassent par E^0,			
largeur encadremt : (c-b)/c	7.88%	21.1%	très nombreuses, sont du même ordre que			
erreur relative : (b-a)/a	−0.39%	0.4%	celles des séquences de faible longueur; il y a donc explosion du nombre de séquences à explorer.			
c1 : majorant SRI α =8.003e-7	3.650e-8	*3.650e-7*	*3.650e-6*	*3.650e-5*	*3.649e-4*	*3.643e-3*
b1 : minorant SRI (estimation)	0	0	*2.193e-6*	*2.863e-5*	*2.930e-4*	*2.930e-3*
temps de réponse	9 secondes pour trouver les 102 séquences de probabilité asymptotique > 2.2e-12 qui vont à la panne sans repasser par E^0					
largeur encadremt : (c1-b1)/c1	100%	100%	39.9%	21.6%	19.7%	19.6%
erreur relative : (c1-a)/a	58.8%	22.9%	19.1%	19%	19%	18.9%